


# Supplementary Materials: User Trust Inference in Online Social Network: A Message Passing Perspective

Yu Liu  and Bai Wang

## 1. Feature statistics

As listed in Table 2 in the paper, the types of features are user profile statistical features ( $\mathcal{F}_{\text{PRF}}$ ), user review linguistic and stylistic features ( $\mathcal{F}_{\text{UGC}}$ ), propagative trust features ( $\mathcal{F}_{\text{TP}}$ ) and auxiliary features. Tables 3, 4 and 5 give brief descriptions for each feature from the former three types about how they are built. Based on the Ising / Potts model, the auxiliary features are in two categories: one ( $\mathcal{F}_{\text{TAUX}}$ ) is for any edge between a user node and a trustRelation node, the other ( $\mathcal{F}_{\text{TPAUX}}$ ) is for edges between trustRelation nodes. Here, we further provide detailed statistics on features used in the paper.

For user profile statistical features in feature set  $\mathcal{F}_{\text{PRF}}$  and user review linguistic and stylistic features in feature set  $\mathcal{F}_{\text{UGC}}$ , all of their values are uniformly scaled to be integers in  $[0, 100]$ . Through such scaling, each of these features is guaranteed to be discrete. For example, features named with “nReviews” prefix will have a total of 101 features of names “nReviews0”, “nReviews1”, ..., “nReviews100”. Furthermore, each feature will be attached to either a node with a specific state or an edge with a specific state-state pair, and then the total number of features of a same name will be multiplied by the number of states of the node or the number of state-state pairs of the edge. In this way, the numbers of node and edge features used for the proposed model are listed in Table S1. Besides, although each edge feature has a corresponding mirrored feature in our implementation of the proposed model, due to the way we implement an undirected edge as two opposite directed edges, the feature and its mirrored feature serve a same purpose, they are counted as one feature in the table.

**Table S1.** Feature statistics for the proposed model.

Feature set	Feature affinity	Count	Note
$\mathcal{F}_{\text{PRF}}$	Node or Edge	1,010	E.g., nReviews, nTrustors, ...
$\mathcal{F}_{\text{UGC}}$	Edge	1,515	E.g., rNouns, rPositives, rHedges, ...
$\mathcal{F}_{\text{TP}}$	Node	16	E.g., d000, d001, ..., d111, ..., t111
$\mathcal{F}_{\text{TAUX}}$	Edge	12	6 for $E_{u \leftrightarrow t}$ , 6 for $E_{t \leftrightarrow u}$
$\mathcal{F}_{\text{TPAUX}}$	Edge	2	For $E_{t \leftrightarrow t}$
Node features		3,062	$= 1010 \times 3 + 16 \times 2$
Edge features		15,230	$= (1010 + 1515 + 12) \times 6 + 2 \times 4$
All features		18,292	$= 3062 + 15230$

## 2. Full results

In this section, the full results for all experiments from the first set and the second set are listed.

### 2.1. Results for the first experiment set

In this section, we report results for all eight experiments in the first experiment set. Tables S2, S3, S4 and S5 respectively list the proposed model’s performance results evaluated by Accuracy, Precision, Recall and  $F_1$  score.

**Table S2.** (1<sup>st</sup> set) Experimental result of the proposed model evaluated by Accuracy.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	0.6464	0.7776	0.9678	0.6492	0.9647	0.8886	0.9664	0.9657
60% – 40%	0.6531	0.7960	0.9673	0.6518	0.9672	0.8856	0.9660	0.9658
70% – 30%	0.6496	0.7787	0.9673	0.6550	0.9648	0.8800	0.9675	0.9642
80% – 20%	0.6572	0.7841	0.9684	0.6611	0.9643	0.8805	0.9627	0.9654
90% – 10%	0.6775	0.7795	0.9804	0.6675	0.9706	0.8846	0.9590	0.9600

**Table S3.** (1<sup>st</sup> set) Experimental result of the proposed model evaluated by Precision.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	0.6025	0.8738	1.0000	0.6009	1.0000	0.8554	0.9496	0.9493
60% – 40%	0.6060	0.8621	1.0000	0.6026	1.0000	0.8485	0.9496	0.9496
70% – 30%	0.6052	0.8674	1.0000	0.6051	1.0000	0.8389	0.9496	0.9474
80% – 20%	0.6117	0.8581	1.0000	0.6096	1.0000	0.8445	0.9449	0.9457
90% – 10%	0.6286	0.8709	1.0000	0.6139	1.0000	0.8480	0.9357	0.9382

**Table S4.** (1<sup>st</sup> set) Experimental result of the proposed model evaluated by Recall.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	0.9664	0.6754	0.9389	0.9946	0.9329	0.9488	0.9888	0.9875
60% – 40%	0.9766	0.7294	0.9378	0.9954	0.9377	0.9529	0.9880	0.9874
70% – 30%	0.9628	0.6846	0.9380	0.9936	0.9332	0.9559	0.9908	0.9869
80% – 20%	0.9564	0.7071	0.9401	0.9919	0.9323	0.9478	0.9867	0.9911
90% – 10%	0.9480	0.6825	0.9628	0.9943	0.9442	0.9514	0.9903	0.9893

**Table S5.** (1<sup>st</sup> set) Experimental result of the proposed model evaluated by F<sub>1</sub> score.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	0.7422	0.7619	0.9685	0.7492	0.9653	0.8997	0.9688	0.9680
60% – 40%	0.7479	0.7902	0.9679	0.7508	0.9679	0.8977	0.9684	0.9681
70% – 30%	0.7432	0.7652	0.9680	0.7521	0.9654	0.8935	0.9698	0.9667
80% – 20%	0.7462	0.7753	0.9691	0.7551	0.9650	0.8932	0.9653	0.9679
90% – 10%	0.7559	0.7653	0.9810	0.7591	0.9713	0.8967	0.9622	0.9631

Tables S6, S7, S8 and S9 respectively list SVM's performance results evaluated by Accuracy, Precision, Recall and F<sub>1</sub> score. It's worth noting that Experiment #1 that uses feature set  $\mathcal{F}_{\text{TAUX}}$  isn't available for experiments for all the comparison methods, and the reason is discussed in the paper. Besides, feature set  $\mathcal{F}_{\text{TPAUX}}$  that has been used in experiments #4, #6, #7 and #8 for the proposed model is also unavailable for these methods.

**Table S6.** (1<sup>st</sup> set) Experimental result of SVM evaluated by Accuracy.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	N/A	0.8577	0.7439	0.8193	0.8415	0.9106	0.8339	0.8874
60% – 40%	N/A	0.8589	0.7483	0.8212	0.8437	0.9139	0.8386	0.8924
70% – 30%	N/A	0.8607	0.7509	0.8163	0.8455	0.9120	0.8355	0.8916
80% – 20%	N/A	0.8621	0.7511	0.8062	0.8476	0.9093	0.8313	0.8919
90% – 10%	N/A	0.8627	0.7521	0.8135	0.8518	0.9143	0.8366	0.8972

**Table S7.** (1<sup>st</sup> set) Experimental result of SVM evaluated by Precision.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	N/A	0.8765	0.7625	0.9429	0.8351	0.9074	0.9043	0.8727
60% – 40%	N/A	0.8783	0.7672	0.9460	0.8395	0.9106	0.9061	0.8786
70% – 30%	N/A	0.8813	0.7700	0.9461	0.8426	0.9122	0.9093	0.8810
80% – 20%	N/A	0.8834	0.7710	0.9400	0.8459	0.9127	0.9089	0.8838
90% – 10%	N/A	0.8870	0.7753	0.9348	0.8525	0.9127	0.9125	0.8876

**Table S8.** (1<sup>st</sup> set) Experimental result of SVM evaluated by Recall.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	N/A	0.8497	0.7463	0.6994	0.8710	0.9247	0.7657	0.9206
60% – 40%	N/A	0.8501	0.7497	0.7007	0.8697	0.9277	0.7738	0.9234
70% – 30%	N/A	0.8501	0.7518	0.6906	0.8690	0.9217	0.7640	0.9184
80% – 20%	N/A	0.8505	0.7504	0.6752	0.8691	0.9155	0.7555	0.9150
90% – 10%	N/A	0.8474	0.7455	0.6943	0.8690	0.9260	0.7629	0.9216

**Table S9.** (1<sup>st</sup> set) Experimental result of SVM evaluated by F<sub>1</sub> score.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	N/A	0.8629	0.7543	0.8031	0.8527	0.9160	0.8293	0.8960
60% – 40%	N/A	0.8639	0.7584	0.8051	0.8543	0.9190	0.8347	0.9004
70% – 30%	N/A	0.8654	0.7608	0.7984	0.8556	0.9170	0.8303	0.8993
80% – 20%	N/A	0.8666	0.7606	0.7859	0.8573	0.9141	0.8252	0.8991
90% – 10%	N/A	0.8667	0.7601	0.7968	0.8607	0.9193	0.8310	0.9043

Tables S10, S11, S12 and S13 respectively list Decision Tree's performance results evaluated by Accuracy, Precision, Recall and F<sub>1</sub> score.

**Table S10.** (1<sup>st</sup> set) Experimental result of Decision Tree evaluated by Accuracy.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	N/A	0.8101	0.6882	0.8193	0.7925	0.8778	0.8303	0.8688
60% – 40%	N/A	0.8124	0.6878	0.8212	0.7904	0.8847	0.8325	0.8780
70% – 30%	N/A	0.8129	0.6918	0.8163	0.7957	0.8853	0.8292	0.8755
80% – 20%	N/A	0.8146	0.6936	0.8062	0.7952	0.8729	0.8263	0.8626
90% – 10%	N/A	0.8159	0.6838	0.8135	0.7954	0.8741	0.8305	0.8689

**Table S11.** (1<sup>st</sup> set) Experimental result of Decision Tree evaluated by Precision.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	N/A	0.8247	0.7312	0.9429	0.8057	0.8778	0.8363	0.8706
60% – 40%	N/A	0.8282	0.7303	0.9460	0.8075	0.8783	0.8369	0.8734
70% – 30%	N/A	0.8273	0.7356	0.9461	0.8100	0.8809	0.8355	0.8737
80% – 20%	N/A	0.8302	0.7372	0.9400	0.8136	0.8793	0.8387	0.8678
90% – 10%	N/A	0.8321	0.7331	0.9348	0.8112	0.8761	0.8373	0.8721

**Table S12.** (1<sup>st</sup> set) Experimental result of Decision Tree evaluated by Recall.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	N/A	0.8122	0.6455	0.6994	0.7988	0.8923	0.8429	0.8821
60% – 40%	N/A	0.8124	0.6459	0.7007	0.7907	0.9068	0.8471	0.8987
70% – 30%	N/A	0.8150	0.6480	0.6906	0.7997	0.9045	0.8415	0.8928
80% – 20%	N/A	0.8146	0.6500	0.6752	0.7929	0.8794	0.8299	0.8719
90% – 10%	N/A	0.8151	0.6285	0.6943	0.7971	0.8865	0.8418	0.8803

**Table S13.** (1<sup>st</sup> set) Experimental result of Decision Tree evaluated by F<sub>1</sub> score.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	N/A	0.8184	0.6856	0.8031	0.8022	0.8850	0.8396	0.8763
60% – 40%	N/A	0.8202	0.6855	0.8051	0.7990	0.8923	0.8420	0.8859
70% – 30%	N/A	0.8211	0.6890	0.7984	0.8048	0.8925	0.8385	0.8831
80% – 20%	N/A	0.8223	0.6909	0.7859	0.8031	0.8793	0.8343	0.8699
90% – 10%	N/A	0.8235	0.6768	0.7968	0.8041	0.8812	0.8396	0.8762

Tables [S14](#), [S15](#), [S16](#) and [S17](#) respectively list Random Forest’s performance results evaluated by Accuracy, Precision, Recall and F<sub>1</sub> score.

**Table S14.** (1<sup>st</sup> set) Experimental result of Random Forest evaluated by Accuracy.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	N/A	0.8573	0.7993	0.8193	0.8634	0.9193	0.8860	0.9212
60% – 40%	N/A	0.8580	0.8013	0.8212	0.8636	0.9215	0.8907	0.9239
70% – 30%	N/A	0.8586	0.8003	0.8163	0.8640	0.9178	0.8863	0.9215
80% – 20%	N/A	0.8585	0.8019	0.8062	0.8655	0.9129	0.8796	0.9152
90% – 10%	N/A	0.8593	0.8055	0.8135	0.8671	0.9152	0.8846	0.9198

**Table S15.** (1<sup>st</sup> set) Experimental result of Random Forest evaluated by Precision.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	N/A	0.8735	0.8333	0.9429	0.8886	0.9250	0.9243	0.9303
60% – 40%	N/A	0.8723	0.8341	0.9460	0.8868	0.9250	0.9262	0.9305
70% – 30%	N/A	0.8724	0.8320	0.9461	0.8871	0.9236	0.9250	0.9314
80% – 20%	N/A	0.8717	0.8347	0.9400	0.8899	0.9196	0.9234	0.9271
90% – 10%	N/A	0.8731	0.8404	0.9348	0.8921	0.9162	0.9234	0.9249

**Table S16.** (1<sup>st</sup> set) Experimental result of Random Forest evaluated by Recall.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	N/A	0.8526	0.7739	0.6994	0.8469	0.9217	0.8536	0.9192
60% – 40%	N/A	0.8557	0.7775	0.7007	0.8496	0.9261	0.8612	0.9246
70% – 30%	N/A	0.8568	0.7781	0.6906	0.8500	0.9200	0.8534	0.9188
80% – 20%	N/A	0.8576	0.7781	0.6752	0.8500	0.9148	0.8412	0.9107
90% – 10%	N/A	0.8575	0.7787	0.6943	0.8506	0.9236	0.8515	0.9228

**Table S17.** (1<sup>st</sup> set) Experimental result of Random Forest evaluated by  $F_1$  score.

Training–Test	# of experiment							
	1	2	3	4	5	6	7	8
50% – 50%	N/A	0.8630	0.8025	0.8031	0.8673	0.9233	0.8875	0.9247
60% – 40%	N/A	0.8639	0.8048	0.8051	0.8678	0.9255	0.8925	0.9275
70% – 30%	N/A	0.8646	0.8042	0.7984	0.8682	0.9218	0.8877	0.9250
80% – 20%	N/A	0.8646	0.8054	0.7859	0.8695	0.9172	0.8804	0.9188
90% – 10%	N/A	0.8653	0.8084	0.7968	0.8709	0.9199	0.8860	0.9238

## 2.2. Results for the second experiment set

In this section, we report results for experiments #2–7 in the second experiment set with reduced feature data. In these experiments, features in the  $\mathcal{F}_{\text{TAUX}}$  set was not involved, and similarly, comparison methods didn't use features from the  $\mathcal{F}_{\text{TPAUX}}$  set. They were conducted on the split 90% – 10% training–train data set with removal of varied percentages of features. The removal ratios were 20%, 40%, 60% and 80%. Results from the first set of experiments with the same setup but no feature removed (the removal ratio is 0%) are also listed for comparison.

Tables S18, S19, S20 and S21 respectively list the proposed model's performance results evaluated by Accuracy, Precision, Recall and  $F_1$  score.

**Table S18.** (2<sup>nd</sup> set) Experimental result of the proposed model evaluated by Accuracy.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.6764	0.7640	0.6668	0.7642	0.6655	0.7609
60%	0.7238	0.8389	0.6695	0.8548	0.7385	0.8047
40%	0.7554	0.8909	0.6719	0.8991	0.7858	0.8650
20%	0.7703	0.9305	0.6692	0.9368	0.8406	0.9153
0%	0.7795	0.9804	0.6675	0.9706	0.8846	0.9590

**Table S19.** (2<sup>nd</sup> set) Experimental result of the proposed model evaluated by Precision.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.6275	0.7073	0.6139	0.7070	0.6125	0.6906
60%	0.6776	0.7932	0.6155	0.8320	0.6721	0.7332
40%	0.7386	0.8578	0.6172	0.9044	0.7195	0.7999
20%	0.7983	0.9207	0.6154	0.9706	0.7844	0.8664
0%	0.8709	1.0000	0.6139	1.0000	0.8480	0.9357

**Table S20.** (2<sup>nd</sup> set) Experimental result of the proposed model evaluated by Recall.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.9493	0.9418	0.9910	0.9433	0.9941	0.9893
60%	0.9074	0.9390	0.9927	0.9076	0.9831	0.9892
40%	0.8292	0.9505	0.9929	0.9041	0.9727	0.9919
20%	0.7546	0.9498	0.9926	0.9074	0.9618	0.9924
0%	0.6825	0.9628	0.9943	0.9442	0.9514	0.9903

**Table S21.** (2<sup>nd</sup> set) Experimental result of the proposed model evaluated by F<sub>1</sub> score.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.7556	0.8079	0.7581	0.8083	0.7580	0.8134
60%	0.7759	0.8599	0.7599	0.8682	0.7984	0.8422
40%	0.7813	0.9018	0.7613	0.9043	0.8271	0.8856
20%	0.7758	0.9350	0.7597	0.9380	0.8641	0.9251
0%	0.7653	0.9810	0.7591	0.9713	0.8967	0.9622

Tables [S22](#), [S23](#), [S24](#) and [S25](#) respectively list SVM's performance results evaluated by Accuracy, Precision, Recall and F<sub>1</sub> score.

**Table S22.** (2<sup>nd</sup> set) Experimental result of SVM evaluated by Accuracy.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.7504	0.5640	0.8135	0.7296	0.8387	0.8320
60%	0.7812	0.6258	0.8135	0.7451	0.8579	0.8409
40%	0.8113	0.6847	0.8135	0.7957	0.8807	0.8573
20%	0.8358	0.7200	0.8135	0.8316	0.8982	0.8672
0%	0.8627	0.7521	0.8135	0.8518	0.9143	0.8366

**Table S23.** (2<sup>nd</sup> set) Experimental result of SVM evaluated by Precision.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.7163	0.7022	0.9348	0.7091	0.9094	0.9331
60%	0.7612	0.6961	0.9348	0.7803	0.8947	0.9170
40%	0.8093	0.7097	0.9348	0.8132	0.8955	0.9019
20%	0.8499	0.7080	0.9348	0.8429	0.9030	0.8821
0%	0.8870	0.7753	0.9348	0.8525	0.9127	0.9125

**Table S24.** (2<sup>nd</sup> set) Experimental result of SVM evaluated by Recall.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.8714	0.2992	0.6943	0.8252	0.7706	0.7337
60%	0.8518	0.5141	0.6943	0.7186	0.8277	0.7675
40%	0.8395	0.6793	0.6943	0.7948	0.8758	0.8180
20%	0.8359	0.7973	0.6943	0.8362	0.9038	0.8633
0%	0.8474	0.7455	0.6943	0.8690	0.9260	0.7629

**Table S25.** (2<sup>nd</sup> set) Experimental result of SVM evaluated by F<sub>1</sub> score.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.7863	0.4196	0.7968	0.7628	0.8343	0.8215
60%	0.8040	0.5914	0.7968	0.7482	0.8599	0.8356
40%	0.8241	0.6942	0.7968	0.8039	0.8855	0.8579
20%	0.8429	0.7500	0.7968	0.8395	0.9034	0.8726
0%	0.8667	0.7601	0.7968	0.8607	0.9193	0.8310

Tables [S26](#), [S27](#), [S28](#) and [S29](#) respectively list Decision Tree's performance results evaluated by Accuracy, Precision, Recall and F<sub>1</sub> score.

**Table S26.** (2<sup>nd</sup> set) Experimental result of Decision Tree evaluated by Accuracy.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.7429	0.5535	0.8135	0.7276	0.8375	0.8146
60%	0.7672	0.5982	0.8135	0.7389	0.8527	0.8061
40%	0.7869	0.6252	0.8135	0.7465	0.8684	0.7907
20%	0.7971	0.6461	0.8135	0.7421	0.8739	0.7697
0%	0.8159	0.6838	0.8135	0.7954	0.8741	0.8305

**Table S27.** (2<sup>nd</sup> set) Experimental result of Decision Tree evaluated by Precision.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.7121	0.7143	0.9348	0.7120	0.9157	0.8750
60%	0.7545	0.6817	0.9348	0.7417	0.8974	0.8294
40%	0.7935	0.6658	0.9348	0.7567	0.8910	0.7845
20%	0.8185	0.6615	0.9348	0.7541	0.8796	0.7440
0%	0.8321	0.7331	0.9348	0.8112	0.8761	0.8373

**Table S28.** (2<sup>nd</sup> set) Experimental result of Decision Tree evaluated by Recall.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.8594	0.2540	0.6943	0.8111	0.7616	0.7560
60%	0.8272	0.4453	0.6943	0.7739	0.8132	0.7956
40%	0.8050	0.5795	0.6943	0.7645	0.8547	0.8310
20%	0.7900	0.6720	0.6943	0.7577	0.8812	0.8582
0%	0.8151	0.6285	0.6943	0.7971	0.8865	0.8418

**Table S29.** (2<sup>nd</sup> set) Experimental result of Decision Tree evaluated by F<sub>1</sub> score.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.7788	0.3747	0.7968	0.7583	0.8316	0.8111
60%	0.7892	0.5387	0.7968	0.7575	0.8533	0.8121
40%	0.7992	0.6197	0.7968	0.7606	0.8725	0.8071
20%	0.8040	0.6667	0.7968	0.7559	0.8804	0.7970
0%	0.8235	0.6768	0.7968	0.8041	0.8812	0.8396

Tables [S30](#), [S31](#), [S32](#) and [S33](#) respectively list Random Forest's performance results evaluated by Accuracy, Precision, Recall and F<sub>1</sub> score.

**Table S30.** (2<sup>nd</sup> set) Experimental result of Random Forest evaluated by Accuracy.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.7470	0.5723	0.8135	0.7558	0.8407	0.7875
60%	0.7779	0.6292	0.8135	0.7802	0.8617	0.8117
40%	0.8058	0.6833	0.8135	0.7945	0.8837	0.8279
20%	0.8304	0.7040	0.8135	0.8036	0.8970	0.8050
0%	0.8593	0.8055	0.8135	0.8671	0.9152	0.8846

**Table S31.** (2<sup>nd</sup> set) Experimental result of Random Forest evaluated by Precision.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.7133	0.7065	0.9348	0.7225	0.9184	0.9230
60%	0.7595	0.6839	0.9348	0.7563	0.9069	0.8807
40%	0.8062	0.6906	0.9348	0.7684	0.9060	0.8245
20%	0.8451	0.6793	0.9348	0.7738	0.9065	0.7676
0%	0.8731	0.8404	0.9348	0.8921	0.9162	0.9234

**Table S32.** (2<sup>nd</sup> set) Experimental result of Random Forest evaluated by Recall.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.8690	0.3218	0.6943	0.8707	0.7655	0.6510
60%	0.8464	0.5507	0.6943	0.8599	0.8219	0.7432
40%	0.8312	0.7225	0.6943	0.8730	0.8694	0.8554
20%	0.8302	0.8300	0.6943	0.8861	0.8971	0.9034
0%	0.8575	0.7787	0.6943	0.8506	0.9236	0.8515

**Table S33.** (2<sup>nd</sup> set) Experimental result of Random Forest evaluated by F<sub>1</sub> score.

% of removed features	# of experiment					
	2	3	4	5	6	7
80%	0.7835	0.4422	0.7968	0.7898	0.8350	0.7635
60%	0.8006	0.6101	0.7968	0.8048	0.8623	0.8061
40%	0.8185	0.7062	0.7968	0.8174	0.8873	0.8396
20%	0.8376	0.7471	0.7968	0.8262	0.9018	0.8300
0%	0.8653	0.8084	0.7968	0.8709	0.9199	0.8860