"Discovering Story Chains:
A Framework Based on Zigzagged Search and News Actors"
Cagri Toraman and Fazli Can

Statistical Tests of Methodology

For Fleiss' Kappa, we use a Java program, which can be downloaded from http://cs.bilkent.edu.tr/~ctoraman/story_chains/FleissKappa.java

Statistical Tests of User Study 1

For the Friedman, Wilcoxon, and Conover tests, the package *PMCMR* in R is used. The details of the Friedman tests are given in Table 2 and 5 in the paper. The details of the Wilcoxon tests are given in Table 3 and 4 in the paper.

The details of the Conover tests are not given in the paper due to the limited space. The p values of the Conover test for Decision 1 are given in Table 10, and for Decision 4 in Table 11.

TABLE 10. The p values of the Conover test for the pairwise comparisons of Figure 6 in the paper.

	Case	*Ukrainian Riots*		*Trucks Going to Syria*		*Allegations to Fenerbahçe*	
Q: *measure*	Algo.	7 days	15 days	7 days	15 days	7 days	15 days
Q1: *relevance*	15 days	0.52		< 0.01		< 0.01	
	30 days	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Q2: *coverage*	15 days	0.05		< 0.01		< 0.01	
	30 days	< 0.01	<0.01	< 0.01	< 0.01	< 0.01	< 0.01
Q3: *coherence*	15 days	-		< 0.01		< 0.01	
	30 days	-	-	< 0.01	< 0.01	< 0.01	0.51
Q4: *disclose relations*	15 days	-		< 0.01		< 0.01	
	30 days	-	-	0.09	< 0.01	< 0.01	0.06

TABLE 11. The p values of the Conover test for the pairwise comparisons of Figure 9 in the paper. VSM: Vector space model, NE: Named entity, SN: Social network.

	Case	*Ukrainian Riots*			*Trucks Going to Syria*			*Allegations to Fenerbahçe*		
Q: *measure*	Algo.	VSM	NE	SN	VSM	NE	SN	VSM	NE	SN
Q1: *relevance*	NE	< 0.01			< 0.01			< 0.01		
	SN	< 0.01	< 0.01		0.527	< 0.01		< 0.01	< 0.01	
	Hybrid	< 0.01	< 0.01	< 0.01	0.016	< 0.01	< 0.01	< 0.01	0.279	< 0.01
Q2: *coverage*	NE	< 0.01			< 0.01			< 0.01		
	SN	< 0.01	< 0.01		0.450	< 0.01		< 0.01	< 0.01	
	Hybrid	< 0.01	< 0.01	< 0.01	0.080	< 0.01	0.270	< 0.01	0.065	< 0.01
Q3: *coherence*	NE	< 0.01			< 0.01			< 0.01		
	SN	< 0.01	< 0.01		0.71	< 0.01		< 0.01	< 0.01	
	Hybrid	< 0.01	< 0.01	< 0.01	0.71	< 0.01	0.59	< 0.01	< 0.01	< 0.01
Q4: *disclose relations*	NE	< 0.01			< 0.01			< 0.01		
	SN	< 0.01	-		< 0.01	< 0.01		< 0.01	< 0.01	
	Hybrid	< 0.01	< 0.01	< 0.01	< 0.01	-	< 0.01	0.013	< 0.01	< 0.01

Statistical Tests of User Study 2

For the Friedman and Conover tests, the package *PMCMR* in R is used. For Cohen's d-test, the package *effsize* in R is used.

The details of the Friedman and Conover tests are given in Table 8 and 9, respectively, in the paper.

The details of Cohen's d-test of User Study 2 are not given in the paper due to the limited space. Cohen's d values and confidence intervals are given in Table 12.

TABLE 12. Cohen's d values and confidence intervals for the pairwise comparison, for Table 9 in the paper.

	Case	*Ukrainian Riots*			*Trucks Going to Syria*			*Allegations to Fenerbahçe*		
Q: *measure*	Algo.	sTDT	aTDT	GN	sTDT	aTDT	GN	sTDT	aTDT	GN
Q1: *relevance*	aTDT	-3.62 [-5.09, -2.14]			1.78 [0.72, 2.84]			0.72 [-0.19, 1.64]		
	GN	-0.77 [-1.69, 0.15]	1.49 [0.48, 2.50]		1.74 [0.69, 2.79]	-0.19 [-1.08, 0.69]		4.03 [2.44, 5.61]	1.81 [0.75, 2.87]	
	hZZ	0.59 [-0.31, 1.50]	6.17 [3.97, 8.37]	1.24 [0.27, 2.22]	2.76 [1.50, 4.03]	0.36 [-0.53, 1.25]	0.71 [-0.20, 1.62]	6.77 [4.39, 9.15]	2.17 [1.04, 3.31]	0.18 [-0.71, 1.06]
Q2: *coverage*	aTDT	-1.85 [-2.93, -0.78]			1.04 [0.09, 1.99]			0.15 [-0.74, 1.03]		
	GN	-0.55 [-1.45, 0.35]	0.88 [-0.05, 1.81]		1.42 [0.42, 2.42]	0.38 [-0.52, 1.27]		1.41 [0.41, 2.41]	1.49 [0.48, 2.50]	
	hZZ	0.50 [-0.40, 1.40]	2.13 [1.01, 3.26]	0.79 [-0.13, 1.71]	1.73 [0.68, 2.78]	0.57 [-0.34, 1.47]	0.12 [-0.76, 1.01]	1.57 [0.55, 2.60]	1.74 [0.69, 2.79]	0.00
Q3: *coherence*	aTDT	-1.30 [-2.28, -0.32]			1.40 [0.40, 2.39]			1.02 [0.07, 1.96]		
	GN	0.00	1.16 [0.20, 2.12]		0.70 [-0.22, 1.61]	-0.68 [-1.60, 0.23]		1.89 [0.81, 2.96]	0.64 [-0.27, 1.54]	
	hZZ	1.21 [0.24, 2.18]	2.95 [1.64, 4.25]	1.08 [0.13, 2.03]	1.77 [0.71, 2.82]	0.07 [-0.82, 0.95]	0.86 [-0.07, 1.79]	1.82 [0.76, 2.89]	0.67 [-0.24, 1.58]	0.07 [-0.82, 0.95]
Q4: *disclose relations*	aTDT	-0.99 [-1.93, -0.05]			0.69 [-0.22, 1.60]			0.33 [-0.56, 1.22]		
	GN	0.00	0.99 [0.05, 1.93]		1.20 [0.23, 2.16]	0.44 [-0.46, 1.33]		1.45 [0.44, 2.45]	0.99 [0.05, 1.93]	
	hZZ	0.00	0.99 [0.05, 1.93]	0.00	1.20 [0.23, 2.16]	0.44 [-0.46, 1.33]	0.00	1.13 [0.17, 2.08]	0.72 [-0.19, 1.64]	-0.24 [-1.13, 0.64]