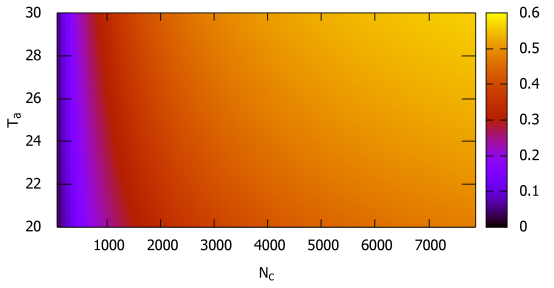


(a) Auditor's reports.

(b) IPO prospectuses.

**Figure 8: The proportion of saved time with respect to $T_a$ and $N_c$.**

to provide related evident paragraphs, and finally give a classification [7, 10, 16]. In academic field, Nuijten et al. [12] propose StatCheck which uses rule-based program to check inconsistency errors in the null-hypothesis significance testing, presented in major psychology journals.

According to the statistics of WebDataCommons, the proportions of entity type, relational type and matrix type of tables in web pages are 59.7%, 38.6% and 1.3%, respectively [9]. However, in vertical domains such as the field of finance, most of tables are matrix-type and have explicit or implicit hierarchical headers. There are some studies about recognizing this type of tables. Fang et al. [3] proposed a Random Forest classification to identify the complex headers in tables; Nagy et al. [11] leveraged rule-based method to extract data categories and data hierarchies from table headers. Based on the extracted tables, there are many understanding tasks, such as linking text to table cells [6], table cell search for a given query [15], ad hoc search over tables [18], transforming complex tables to the form that can be stored in a database [14]. Our task, cross-checking over numerical tables, is also a table understanding task based on extracted table structure.

## 7 CONCLUSION

When investor finds out accounting errors in financial reports, they may doubt firm's governance capacity and authenticity of firm's accounting information. In this paper, we aims at automatic numerical cross-checking over tables in a document, and propose an end-to-end solution to detect whether two table cells are semantically equivalent or not. Based on this model, an auditing tool is built to support guided cross-checking. This system has been widely adopted in the Chinese financial community. Users feedback and theoretical analysis confirm that it helps to save around 50% time.

## 8 ACKNOWLEDGEMENTS

This work was supported by the National Key Research and Development Program of China under Grant No. 2017YFB1002104, the National Natural Science Foundation of China under Grant No. U1811461, and the Innovation Program of Institute of Computing Technology, CAS.

## REFERENCES

[1] Yixuan Cao, Hongwei Li, Ping Luo, and Jiaquan Yao. 2018. Towards Automatic Numerical Cross-Checking: Extracting Formulas from Text. In *WWW*.
[2] Preeti Choudhary, Kenneth J Merkley, and Katherine Schipper. 2019. Do Immaterial Error Corrections Matter? *Available at SSRN 2830676* (2019).
[3] Jing Fang, Prasenjit Mitra, Zhi Tang, and C Lee Giles. 2012. Table header detection and classification. In *AAAI*.
[4] Vivian W Fang, Allen H Huang, and Wenyu Wang. 2017. Imperfect accounting and reporting bias. *Journal of Accounting Research* (2017).
[5] Naeemul Hassan, Fatma Arslan, Chengkai Li, and Mark Tremayne. 2017. Toward Automated Fact-Checking: Detecting Check-worthy Factual Claims by ClaimBuster. In *KDD*.
[6] Dae Hyun Kim, Enamul Hoque, Juho Kim, and Maneesh Agrawala. 2018. Facilitating document reading by linking text and tables. In *UIST*.
[7] Mio Kobayashi, Ai Ishii, Chikara Hoshino, Hiroshi Miyashita, and Takuya Matsuzaki. 2017. Automated Historical Fact-Checking by Passage Retrieval, Word Statistics, and Virtual Question-Answering. In *IJCNLP*.
[8] Alastair Lawrence. 2013. Individual investors and financial disclosure. *Journal of Accounting and Economics* (2013).
[9] Oliver Lehmberg, Dominique Ritze, Robert Meusel, and Christian Bizer. 2016. A large public corpus of web tables containing time and context metadata. In *WWW*. 75–76.
[10] Moin Nadeem, Wei Fang, Brian Xu, Mitra Mohtarami, and James R. Glass. 2019. FAKTA: An Automatic End-to-End Fact Checking System. In *NAACL*.
[11] George Nagy and Sharad Seth. 2016. Table headers: An entrance to the data mine. In *ICPR*.
[12] Michèle B. Nuijten, Chris H. J. Hartgerink, Marcel A. L. M. van Assen, Sacha Epskamp, and Jelte M. Wicherts. 2016. The prevalence of statistical reporting errors in psychology (1985–2013). *Behavior Research Methods* (2016).
[13] Radim Řehůřek and Petr Sojka. 2010. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*.
[14] Alexey O Shigarov, Viacheslav V Paramonov, Polina V Belykh, and Alexander I Bondarev. 2016. Rule-based canonicalization of arbitrary tables in spreadsheets. In *ICIST*.
[15] Huan Sun, Hao Ma, Xiaodong He, Wen-tau Yih, Yu Su, and Xifeng Yan. 2016. Table cell search for question answering. In *WWW*.
[16] James Thorne, Andreas Vlachos, Christos Christodoulopoulos, and Arpit Mittal. 2018. FEVER: a Large-scale Dataset for Fact Extraction and VERification. In *NAACL*.
[17] Andreas Vlachos and Sebastian Riedel. 2014. Fact Checking: Task definition and dataset construction. In *Workshop on Language Technologies and Computational Social Science, ACL*.
[18] Shuo Zhang and Krisztian Balog. 2018. Ad hoc table retrieval using semantic similarity. In *WWW*.