Resolving Ambiguities in Regulations

Towards Achieving the Kohlbergian Stage of Principled Morality

Smita Ghaisas
Tata Research Development and Design Center (TCS Research), Pune, India
smita.ghaisas@tcs.com

Abhishek Sainani
Tata Research Development and Design Center (TCS Research), Pune, India
a.sainani@tcs.com

Preethu Rose Anish
Tata Research Development and Design Center (TCS Research), Pune, India
preethu.rose@tcs.com

ABSTRACT

According to Kohlberg, the final stage of morality is characterized by viewing laws as a means to an end by upholding values such as human dignity and fairness as guiding principles for complying with the essence of the law. Given that purpose of compliance is indeed wellbeing of citizens, software systems should, by design, incorporate these values so that laws are followed in spirit. How can we build software systems that incorporate these values? We present our work on disambiguating Health Insurance Portability and Accountability Act (HIPAA) so as to reduce the potential incidents of breach, thereby upholding of the aforesaid guiding principles of morality. We have employed deep learning based approaches to emulate the human process of disambiguation by integrating information from multiple sources, summarizing it, and augmenting the regulatory text with the additional information. This augmented regulatory text can be used by policy makers and software engineers to achieve compliance in spirit.

CCS CONCEPTS

• Security and privacy~Social aspects of security and privacy • Security and privacy~Privacy protections • Social and professional topics~Patient privacy • Computing methodologies~Information extraction

KEYWORDS

Regulations, Ambiguity Resolution, Principled Morality, Deep Learning, Compliance

1 INTRODUCTION

Regulations are meant to serve wellbeing of citizens. It is therefore desirable that regulatory compliance is guided by social values such as human dignity and fairness. According to Kohlberg [20], the pre-conventional morality stage of human beings is characterized by obedience or suppressing one's desires only because of fear of punishment or for self-interest whereas the conventional morality stage involves rote adherence to laws, or social conventions, to maintain order. Some of the recent works [21, 23], on ensuring compliance seem to reflect these Kohlbergian stages. The final post-conventional stage, also called the stage of principled morality, however, involves viewing laws as a means to an end and puts values such as human dignity and fairness as guiding principles for complying with the essence of the law. We posit that software systems should, by design, incorporate these guiding principles so that laws are followed in spirit and not just in letter. How can we build systems that incorporate these values?

Since regulations are ambiguous, the process of deriving system requirements is ad hoc and error prone [1, 12]. By virtue of these ambiguities in regulations, organizations that must adhere to them often fail to prevent breaches that happen in real life and regulations are increasingly seen to be subject to misuse, abuse and violation [1]. If the potential breaches are arrested by way of disambiguation of the text, we may be able to build systems that inherently value and enable human dignity and fairness. In highly regulated and ubiquitous domains such as healthcare, this becomes all the more critical. To demonstrate this, we take the US Health Insurance Portability and Accountability Act (HIPAA) as an example. As of late May 2017, the, the Office of Civil Rights (OCR) collected close to $15 million dollars in HIPAA violation settlements [3]. Such violations cost huge monetary damage to the healthcare organization, but a more serious concern is the violation of human dignity and fairness [20] of an individual. It is not hard to imagine that most individuals, upon disclosure of sensitive medical information such as HIV status, reproductive issues, certain genetic markers and diagnoses such as cancer and mental health issues would experience a devastating impact on their lives [4]. This can lead to long lasting embarrassment, mental distress, damage to relationships, or social ostracism [4, 8].

Despite HIPAA setting forth certain requirements for safeguarding a patient's privacy and security, many breaches keep happening on a regular basis [22]. To understand the reasons for these breaches, we analyzed published HIPAA case studies on internet. A HIPAA breach is either due to an accidental misuse or

ICSE-SEIS'18, May 27-June 3, 2018, Gothenburg, Sweden
© 2018 Association for Computing Machinery.
ACM ISBN 978-1-4503-5661-9/18/05...$15.00
https://doi.org/10.1145/3183428.3183433

a malicious misuse. Malicious misuses correspond to outsider attacks, whereas accidental misuses correspond to insider attacks or human errors [5]. Our research focuses on the accidental misuse.

We found the following main reasons for accidental misuse of HIPAA clauses [9, 10]:

- Lack of employee training.
- Poorly implemented physical, technical and administrative safeguards.
- Limited understanding of consent and authorization based HIPAA regulations.

We observed that all of these could be attributed to the ambiguous nature of HIPAA regulations. Ambiguity arises when a statement is missing relevant information or when a word or phrase has more than one possible interpretation [11]. For instance, consider the following access related breach:
In 2010, Triple-S Management Corporation (TSMC) reported to OCR that it discovered that two of its former workforce members currently employed by a competitor improperly accessed restricted areas of TSMC's proprietary internet IPA database. They were able to gain access to the database because their access rights were not terminated upon leaving the employment of TSMC. The electronic Protected Health Information (ePHI) accessed in the database included members' names, contract numbers, home addresses, diagnostic codes and treatment codes. TSMC had to pay $3.5 Million and as corrective action plan was asked to adopt a robust plan to correct deficiencies in its HIPAA compliance program [6].

The corresponding HIPAA clause [2] on Access Control (§ 164.308(a)(4)) states: "*Implement technical policies and procedures for electronic information systems that maintain ePHI to allow access only to those persons or software programs that have been granted access rights*." However, the term "access" in this clause and the sub clauses therein is ambiguous.

An online document from the National Institute of Standards and Technology (NIST) [7] in its Introductory Resource Guide for Implementing HIPAA Security Rule states the following for establishing termination procedure: "*Implement procedures for terminating access to ePHI when the employment of a workforce member ends or as required by determinations made as specified in §164.308(a)(3)(ii)(B). Develop a standard set of procedures that should be followed to recover access control devices (Identification [ID] badges, keys, access cards, etc.) when employment ends. Deactivate computer access accounts (e.g., disable user IDs and passwords).*" Had this additional information source been taken into account, this breach could have been avoided. Human beings usually look at multiple sources to make sense of texts that challenge their comprehension. The disambiguation is a cumulative outcome of the integration of knowledge from multiple sources such as official government websites, university websites, FAQs, Forums, Press release etc.

In this paper, we present our ongoing work on an automated approach to emulate this human process by integrating additional information from multiple sources, summarizing it, and augmenting the regulatory text with additional information. This augmented regulatory text can be used by policy makers and software engineers to achieve compliance in spirit.

2 AMBIGUITY RESOLUTION FRAMEWORK

Fig. 1 presents a schematic of the ambiguity resolution framework. When a person interprets a regulation statement, he tries to decipher the meaning of a given term, a phrase or the whole statement within a given context or domain. For this, the person may need to refer to additional source(s) of information to help him map the high level concepts of regulations to the low level implementation details such that the regulation is enforced effectively. This much needed mapping helps disambiguate a regulation and amounts to establishing traceability between a regulation statement and its corresponding relevant additional information. Such additional information are of two types:

- **Internal:** Information within the regulatory document that can alter the meaning of a regulation. This includes information pertinent to References, Amendments, Applicability and Definitions.
- **External:** Information present in external sources such as official government websites (HHS), university websites, FAQs, Forums, Resolution Agreements, Press release of HIPAA settlement cases and judgments.

The challenge we face when automating this human process is the negligible similarity between the regulation and its corresponding additional relevant information due to term mismatch (as is evident from our example regulation and its corresponding additional relevant information discussed in Introduction section).

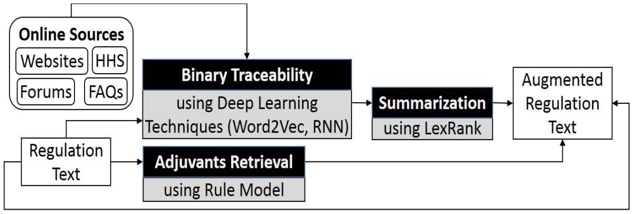

Figure 1: Ambiguity Resolution Framework

Traditional feature-based machine learning techniques fail to identify and retrieve relevant statements from online sources due to such term mismatch between source and target artefacts [13]. To address this problem, we have employed a combination of different techniques for identifying the internal and external additional information. For identifying internal information, we used a variant of Rule Model reported by [19]. For identifying external information, which is considerably larger in size in comparison to internal information, we used Word Embedding to learn semantic information of the domain vocabulary and Recurrent Neural Networks (RNNs) [17] to capture sequential information of the text data. Since the degree of redundancy in relevant information from multiple sources is much higher than degree of redundancy in information from a single source, we use a summarization technique called LexRank [14] to reduce this redundancy. We augment the regulation with the internal information and summarized external information.

For our experiments, we have considered following major clauses in HIPAA security standard: Administrative safeguards (45 CFR Part 164.308), Physical safeguards (45 CFR Part 164.310), and Technical safeguards (45 CFR Part 164.312). Even though the regulation document provides "implementation specifications", the details provided are severely limited to be directly applied. For instance, consider the implementation specification *Encryption and decryption* that states, "*Implement a mechanism to encrypt and decrypt ePHI*." This specification clearly needs more details pertinent to how encryption and decryption has to be done, for instance, whether it has to be done during storage, or transmission of the ePHI. Thus, including relevant additional information would help disambiguate such regulations and contribute to their implementation in spirit.

2.1 Identifying internal additional information

We identified the following four types of internal information that is essential to understand a regulation completely:

Amendment: This includes the information related to modifications or updates that need to be considered for interpreting a regulation within a given context.

Applicability: This includes information that mentions the criteria or scope and time when the regulation comes into effect.

Definition: This includes the information that explains the meaning of a term or phrase used in regulations.

Reference: This includes the information that directs reader to the sources of additional information such as other paragraphs within the document.

To automatically detect the internal information pertinent to a given regulation, we created patterns using a combination of Parts of Speech (POS) tags, keywords and wild card characters [19], upon manually analysing the three major clauses in HIPAA security standard mentioned in section 2.

We created a total of 11 such patterns which included 3 of type *Definition*, 1 of type *Amendment*, and 7 of type *Reference*. Applicability is present as a separate subsection, hence we directly map the regulation statement with its corresponding applicability subsection.

2.2 Identifying external additional information

In this section, we describe the process for identifying additional information from external sources.

2.2.1 Data collection and preparation. In this step, we collected HIPAA related data from various online sources such as websites, discussion forums, blogs, audit sites, and court settlements. The data was in different file formats: pdf, docx, doc, and scanned images. We converted the entire dataset into raw text form of size 15 MB. In preprocessing, we converted the raw text to lowercase and removed non-alphanumeric characters except hyphen and underscore.

We divided our preprocessed dataset into two sets: source artifacts and target artifacts. Source artifact consisted of regulation statements and target artifact consisted of HIPAA related additional data collected from online sources.

2.2.2 Binary Traceability. This step is divided in two phases:

Training Phase: In this phase, we trained the model to identify relevant information. For this purpose we used 45% of the dataset. The training phase was as follows:

1. We used a Neural Networks based Word Embedding technique called Word2Vec [16] to map each term present in the complete dataset, to their corresponding vector form in vector space based on their semantic position in the dataset.
2. We then used a single layer RNN that takes as input vector form of terms (based on Word2Vec model) sequentially for a given sentence and gives a single vector of fixed length as output. We used RNN on sentences of source artifacts and target artifacts to produce a vector output of fixed length for both source artifacts and target artifacts.
3. The similarity score between the two vectors was calculated using Cosine Similarity [18]. If the score crossed a desired threshold then we considered the statements in target artifact as relevant to source artifact. We decided on the desired threshold as 0.65 at which we got highest value of F-score.
4. To train our model, we used a negative log likelihood as a loss function to be minimized. Based on this function we used a stochastic gradient descent method to update the RNN parameters. We tuned the learning rate, number of epocs and other parameters in gradient descent method to optimally train the model.

Testing and Validation Phase: We ran the trained model on the remaining 10% dataset for validation and 45% dataset for testing as follows:

1. We used Word2Vec model to get the semantic vector values for each term in a given input statement.
2. Then we gave these word vectors sequentially as input to RNN. The output of RNN, a single fixed length vector, was obtained for source artifact as well as for target artifact.
3. The similarity between the vectors was calculated using Cosine Similarity. If the score crossed a desired threshold then we considered the statements in target artifact as relevant to source artifact.

2.2.3 Result and Discussions. We ran the model for different variants of Word2Vec and RNN and observed that Skip-gram variant of Word2Vec and Bi-GRU variant of RNN gave best results. We obtained a precision of 22% on recall of 60%. This is based on the true links we obtained from various online sources such as website of New York University [26], where they have provided their policies and procedures corresponding to HIPAA regulations. Due to a skewed dataset (only 55 regulation statements in source dataset whereas 437 statements in target dataset), we obtained a low precision. A single regulation statement was getting mapped with many different statements from online sources. The dataset for traceability also suffered from class imbalance, i.e., out of 24,035 total traceability links, only 1.8% links were true links. This led to lower recall in our results. Our dataset size is nearly 15 MB and contains nearly 42,000 vocabulary terms of pure text relevant to HIPAA regulations. Since collecting relevant information and creating true links for training traceability model was a manual task, the dataset size has been a constraint in our work.

2.2.4 Summarization. The external additional information collected based on Binary Traceability technique was summarized using state-of-the-art summarization technique called LexRank and its implementation in Sumy, a Python Library [15]. We obtained an average precision of 70% and recall of 40%.

Table 1 shows an example of the final output of our framework, i.e., an augmented regulation.

Table 1: Example Augmented Output

Regulation	Internal information	External information
Encryption and decryption. Implement a mechanism to encrypt and decrypt ePHI.	Definition of encryption as mentioned in the HIPAA Security Standards section.	• Possible scenarios of use of encryption and decryption, such as, transmission over network, storage, and sending an email containing ePHI. • Required possible technical details. • Mechanism details for storing the cryptographic keys. • Settlement or court cases regarding HIPAA violation due to lack of encryption.

4 CONCLUSION AND FUTURE WORK

We discuss an automated approach to disambiguate a given regulatory text by integrating additional information from multiple sources, summarizing it, and augmenting the regulatory text with the additional information. We employ deep learning based approaches to facilitate this disambiguation. We believe this to be a step towards the Kohlbergian stage of principled morality wherein social values are as the guiding principles for complying with the essence of the law.

In our experiments with establishing traceability between a regulation statement and its corresponding additional information, we obtained a precision of 22% on a recall of 60%. These are early results. We are exploring technical refinements to improve our results such as use of Transfer Learning [25] in enhancing Word2Vec model and use of Siamese architecture [24] as an alternative to a single RNN. We are also looking into enhancing our dataset to make it bigger and more balanced by developing a semantic crawler based on enhanced Word2Vec model.

REFERENCES

[1] T. D. Breaux, M. W. Vail, and A. I. Anton, A.I., 2006. Towards regulatory compliance: Extracting rights and obligations to align requirements with regulations. In *Requirements Engineering, 14th IEEE International Conference* (pp. 49-58). IEEE.

[2] Health Insurance Portability and Accountability Act, USC H.R. 3103-168, April 2000.

[3] 5 Lessons Learned from 2016-2017 OCR HIPAA Settlements. https://www.paubox.com/blog/lessons-learned-ocr-hipaa-settlements Last accessed on 08-02-2018

[4] The nature of harm in a data breach. https://www2.idexpertscorp.com/knowledge-center//single/the-nature-of-harm Last Accessed on 08-02-2018.

[5] Ö. Kafali, J. Jones, M. Petruso, L. Williams, M. P. Singh, 2017, May. How good is a security policy against real breaches? A HIPAA case study. In *Proceedings of the 39th International Conference on Software Engineering* (pp. 530-540). IEEE Press.

[6] Triple-S Management Corporation Settles HHS Charges by Agreeing to $3.5 Million HIPAA Settlement. https://www.hhs.gov/hipaa/for-professionals/compliance-enforcement/agreements/triple-s-management/index.html Last accessed on 08-02-2018.

[7] http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-66r1.pdf Last accessed on 08-02-2018

[8] Jack Brill, Giving HIPAA Enforcement Room to Grow: Why There Should Not (Yet) Be a Private Cause of Action, 83 Notre Dame L. Rev. 2105 (2008).

[9] Jackson, J., 2015. The Costs of Medical Privacy Breach. MD advisor: a journal for New Jersey medical community, 8(3), pp.4-12.

[10] R. Hsieh, 2014. Improving HIPAA Enforcement and Protecting Patient Privacy in a Digital Healthcare Environment. *Loy. U. Chi. LJ*, *46*, p.175.

[11] J. R. Reidenberg, J. Bhatia, T. D. Breaux, and T. B. Norton, 2016. Ambiguity in privacy policies and the impact of regulation. *The Journal of Legal Studies*, *45*(S2), pp.S163-S190.

[12] A. K. Massey, R. L. Rutledge, A.I. Antón, P. P. Swire, Identifying and classifying ambiguity for regulatory requirements, in 22nd IEEE International Requirements Engineering Conference (RE), Karlskrona, Sweden, 2014, pp. 83–92.

[13] J. Guo, J. Cheng, and J. Cleland-Huang. Semantically enhanced software traceability using deep learning techniques. In Proceedings of the International Conference on Software Engineering, ICSE '17

[14] G. Erkan and D.R. Radev, "LexRank: Graph-based lexical centrality as salience in text summarization," Journal of Artificial Intelligence Research, vol. 22, pp. 457-479, 2004.

[15] Sumy, a Python library and command line utility version 0.7.0. https://pypi.python.org/pypi/sumy, Last accessed on 08-02-2018.

[16] Y. Goldberg and O. Levy (2014). word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method. arXiv:1402.3722 [cs, stat]

[17] T. Mikolov, M. Karafi´at, L. Burget, J. Cernock`y, and S. Khudanpur. Recurrent neural network based language model. In Interspeech, volume 2, page 3, 2010.

[18] J. Human Hayes, A. Dekhtyar, and S. K. Sundaram. Advancing candidate link generation for requirements tracing: The study of methods. IEEE Transactions on Software Engineering, 32(1):4–19, 2006.

[19] S. Ghaisas, M. Motwani, P. R. Anish, Detecting system use cases and validations from documents, In proceedings of the 24th IEEE/ACM Intl. Conference on Automated Software Engineering (2013), pp.568-573.

[20] L. Kohlberg (1969). The cognitive development approach to socialization in D. Goslin (Ed.) Handbook of Socialization Theory and Research (pp. 347-480). Chicago: rand McNally.

[21] S. Ghaisas, A. Sainani, P. R. Anish, R. Suriyanarayanan, and P. Rajaram, 2017, May. Ethos, pathos, and logos to prevent sexual harassment at workplaces: a regulatory solution based on operant conditioning. In *Proceedings of the 39th International Conference on Software Engineering Companion* (pp. 222-224). IEEE Press.

[22] K. Colorafi, and B. Bailey, 2016. It's Time for Innovation in the Health Insurance Portability and Accountability Act (HIPAA). *JMIR medical informatics*, *4*(4).

[23] T. Herath, and H. R. Rao, 2009. Protection motivation and deterrence: a framework for security policy compliance in organisations. *European Journal of Information Systems*, *18*(2), pp.106-125.

[24] J. Mueller and A. Thyagarajan, "Siamese Recurrent Architectures for Learning Sentence Similarity," in Proc. AAAI, 2016.

[25] S. Pan and Q. Yang. A survey on transfer learning. Knowledge and Data Engineering, IEEE Transactions on, 22(10):1345–1359, 2010.

[26] NYU HIPAA Policies, https://www.nyu.edu/about/policies-guidelines-compliance/policies-and-guidelines/hipaa-policies.html, last accessed on 08-02-2018.