RESEARCH FEATURE

Privacy-Knowledge Modeling for the Internet of Things: A Look Back

Charith Perera, Open University
Chang Liu, CSIRO
Rajiv Ranjan, China University of Geosciences and Newcastle University
Lizhe Wang, China University of Geosciences
Albert Y. Zomaya, University of Sydney

Together, the Internet of Things (IoT) and cloud computing give us the ability to gather, process, and even trade data to better understand users' behaviors, habits, and preferences. However, future IoT applications must address the significant potential threats to privacy posed by such knowledge-discovery activities.

The Internet of Things (IoT) connects people and things anytime, anyplace, and ideally using any path, network, and service.[1–3] Over the past few years, numerous IoT solutions have reached the marketplace.[4] Together, these solutions collect a significant amount of data that can be used to derive useful, but extremely personal, knowledge about users.[5] At the same time, cloud computing provides ubiquitous, convenient, on-demand access to a shared pool of configurable computing resources.

We developed the open data market business model to derive value out of such data.[5] This type of market envisions the exchange of data between different parties in the IoT domain. Data owners will collect data using IoT products and solutions; interested consumers will then negotiate with the data owners to obtain access to the data; and data captured by IoT products will help consumers to understand the preferences and behaviors of data owners and to generate additional business value using techniques ranging from waste reduction to personalized service offerings. In open data markets, data consumers will be able to give back part of the additional value generated to the data owners.

However, as IoT services become more powerful and cheaper and as open data markets facilitate data trading, the risk of user privacy violations increases significantly. Therefore, it is important to define privacy explicitly and understand what privacy means for each user of a given system to ensure that personal information is protected at all times.

The late Alan F. Westin, credited with creating the modern field of privacy law, defined *information privacy* as "the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others."[6] *Knowledge modeling* is the process of creating a computer-interpretable model of knowledge or standard specifications about a process, product, or concept. In this article, we consider any piece of information that can be used to understand an individual's privacy expectation, in any given context, to be *privacy knowledge*.

0018-9162/16/$33.00 © 2016 IEEE

TABLE 1. Platform for Privacy Preferences (P3P) protocol privacy policy.

Information stored by the server	Use of collected information	Permanence and visibility
Which kind of information is collected (identifying or not)	How this information is used (for regular navigation, tracking, personalization, telemarketing)	How long information is stored
Which particular information is collected (IP address, email, name)	Who will receive this information (current company or third party)	Whether and how the user can access the stored information (read-only, opt in, opt out)

Our objective in this article is to survey how privacy knowledge has been modeled in the past in different domains, specifically looking at major efforts that attempted to give privacy control to users. Although this article is not an exhaustive review of past work, we hope to capture insights from a broad range of approaches by analyzing different privacy modeling approaches to identify any common patterns and applications. We specifically discuss how these past approaches are relevant in the IoT domain. Unlike the web domain, which only captures users' online activities, IoT systems can capture users' activities and behaviors 24/7—online and offline—through a variety of devices. Because of this accessibility, however, the IoT domain poses more significant privacy risks.

The ultimate goal of this survey is to provide insights into future work, especially identifying major research challenges such as the importance of developing a comprehensive privacy-knowledge model for the IoT and of developing nonintrusive user privacy-preference knowledge-acquisition techniques.

PRIVACY-KNOWLEDGE MODELING: HISTORICAL VIEW

One of the major privacy-preference modeling approaches of the past is the Platform for Privacy Preferences (P3P; w3.org/P3P).[7] P3P is a key milestone in efforts to model privacy preferences, despite its limitations and lack of adoption, which we will discuss shortly. Because it was designed for the web domain (not the IoT), P3P can only attempt to protect user privacy during online browsing activities. To propose the next-generation privacy-preference modeling approaches, especially for newer paradigms such as open data markets in the IoT, it is important to understand what P3P is, how it was designed to work, and why it failed.

The initial intention of P3P was to give users more control over their personal information while web browsing. P3P is an XML-based machine-readable language that helps express a website's data-management practices. In P3P, information is managed based on the users' privacy preferences and the websites' privacy policies.

P3P works as follows. First, websites specify a set of policies that defines how the site intends to use the personal information that might be gathered from visitors. On the other end, users are required to define their own set of preferences for the collection and processing of personal information by the sites they visit. When a user visits a site, P3P compares the user's policy with that of the website. The comparison primarily evaluates what personal information the user is willing to release and what information the website wants to receive. If the two sets of privacy policies do not match, P3P will advise users and ask if they are willing to proceed to the site despite the risk of giving up more personal information.

When using the P3P policy, however, users are in a somewhat helpless situation with limited options. P3P was designed as a way to express privacy preferences but not as a negotiation framework. Therefore, P3P privacy profiles are static. For example, a user might define a privacy policy saying that information about his or her browsing habits should not be collected. If a website's policy states that it uses a cookie for this purpose, the browser automatically rejects the cookie. However, it is likely that key parts of the website's functionality will depend on that cookie, so if it is rejected, the user's experience will be degraded. As a result, most of the time, websites tend to get the information they want. The only exception would be if large numbers of users decided not to visit a particular website because of its unreasonable privacy policies. In that case, the website could be pressured implicitly to change its policies. Table 1 summarizes the main elements of a privacy policy in P3P.

Ultimately, P3P suffered from limited adoption. Despite its potential benefits, P3P failed to receive the necessary attention from browser makers, Internet advocates, and institutions.[8] The Electronic Privacy Information Center (EPIC) described P3P as a "complex and confusing protocol that will make it more difficult for Internet users to protect their privacy."[8] One of P3P's drawbacks is that it will effectively exclude good websites that lack P3P code, even though the privacy practices of these sites might far exceed P3P-compliant sites. Another challenge is the lack of any means to enforce privacy policies. An additional criticism of P3P is over a lack of effective ways to educate users about the levels of privacy and what P3P actually does to protect them. Without this information, it is hard for nontechnical users to understand and configure P3P based on privacy expectations.

Nevertheless, P3P is important to consider because it allows users to define their privacy expectations, putting them in control of their privacy.

Machine interpretability is one of the key objectives in P3P. For example, once both websites and users define their privacy preferences, the remaining interactions occur almost autonomously with minimal human intervention. Because we envision a large number of data-trading transactions in open data markets, this is also one of the main priorities to consider.

In open data markets, data owners and consumers should be able to collectively define their privacy preferences and policies in such a way that machines can take over the trading negotiations and act on behalf of both data owners and consumers. Building a common knowledge model using ontologies increases the ability to conduct trading activities autonomously. P3P intended to create an open and transparent method to express privacy preferences and to make it easier for users to set privacy policies.

PRIVACY MODELING IN OPEN DATA MARKETS

Here, we discuss why we need something similar to P3P in open data markets in the IoT domain for data trading to occur, and why such an approach would work despite P3P's failure. It is also important to identify ways to overcome the issues that disrupted P3P's adoption. In open data market environments, each data owner can have his or her own privacy preferences. The data users would like to trade with other parties might also vary based on a number of factors, such as the type of data, type of data consumer, purpose of data collection, how the data is managed, risks involved, expected return, and so on. From the data consumers' point of view, factors such as data accuracy, data capture frequency, data communication frequency, and the amount of value that can be generated by using the data play a significant role when deciding whether to buy data from a certain data owner.

Data consumers generally are not interested in buying data from one or two individual data owners. To derive useful knowledge, data consumers need to gather and process data on a large scale (more than 10,000 users). Such knowledge will help data consumers reduce waste or generate new customer value. For example, a supermarket chain (a data owner) might use the data collected to optimally manage their supply chain and effectively reduce waste in consumer goods. Part of such value will be given back to data owners to attract them again as potential data sellers. In traditional market settings, a trade would occur when the buyer perceives a certain product (or service) as equally or more valuable than the price the seller is expecting in return. Similarly, a data trade between a data owner and consumer depends on perceived privacy risks and benefits. If the data owners perceive that they will get a return sufficient to trade off the privacy risks involved, they will agree to sell their data.

In open data markets, we intend to use privacy-preference models to conduct data-trading negotiations between data owners and consumers instead of making strict decisions based on static privacy profiles. That means both data owners and consumers might be willing to change their privacy expectations based on privacy risks and rewards involved in each trading occurrence. To conduct data trading as well as perform automated risk-benefit negotiations, we need to capture and model a certain amount of information from both data owners and consumers.

To support P3P, the World Wide Web Consortium (W3C) developed and recommended a language called APPEL (A Privacy Preference Exchange Language; w3.org/TR/P3P-preferences) that can be used to express user privacy preferences. With APPEL, users can express their privacy preferences through a set of preference rules, which can then be used by software to make automated or semiautomated decisions regarding the acceptability of machine-readable privacy policies from P3P-enabled websites. The major weakness of APPEL is that it can only specify what is unacceptable, but not what is acceptable, for a user. The IoT demands a more comprehensive approach toward privacy modeling because the goal is for it to deal with both our online and offline everyday lives. Because of this, privacy risks will grow exponentially.

EVALUATION OF RESEARCH EFFORTS: MODELING AND APPLICATIONS

Here, we review a range of past approaches to modeling privacy knowledge in different application contexts. Table 2 summarizes the following discussion by listing each approach, its primary application domain, and the factors included in its knowledge models.

Nonontology-based privacy-knowledge modeling

Zakwan Jaroucheh and his colleagues proposed a context information dissemination framework based on privacy policies.[9] Their application domain is a smart university, where staff and students at different universities can

TABLE 2. Summary of factors modeled in past privacy-knowledge modeling approaches.

Modeling approach	Modeling language	Primary application domain	Factors identified and modeled														
			Data	Data consumer	Retention period	Purpose	Dispute/remedy	Access	Policy/disclosure	Actions	Obligations/limits	Data consumer (opt out)	Purpose (opt out)	Consent	Techniques used*	Incentive	Trust
Jaroucheh[a]	XML	Smart university	✗	✗	✗												
Zhang and Todd[b]	Ontology	Ubiquitous computing	✗	✗		✗	✗	✗									
Sacco and Passant[c]	Ontology	Linked data	✗	✗		✗		✗									
Hu and Yang[d]	Ontology	Healthcare	✗	✗	✗	✗		✗	✗	✗	✗	✗	✗				
Kost and Freytag[e]	Ontology	Transportation	✗	✗	✗	✗		✗	✗		✗			✗			
Ahmed[f]	Ontology	Personal information system	✗	✗	✗	✗			✗						CTS		
Panagiotopoulos[g]	Ontology	Mobile e-commerce	✗	✗	✗											✗	
Youssef[h]	Ontology	–	✗	✗	✗	✗		✗									
Bodorik[i]	Ontology	Web services	✗	✗	✗	✗	✗	✗		✗	✗						
Garcia[j]	Ontology	E-commerce	✗	✗	✗	✗			✗					✗			
Kagal[k]	Ontology	Ubiquitous computing	✗	✗		✗		✗									✗
Hecker[l]	Ontology	Service-oriented architecture	✗	✗	✗	✗		✗	✗	✗	✗			✗	C		
Martimiano[m]	Ontology	Ubiquitous computing	✗	✗	✗	✗		✗	✗	✗					CTS		✗

* Data collection (C), data transmission (T), and data storage (S)

(a) Z. Jaroucheh, X. Liu, and S. Smith, "An Approach to Domain-Based Scalable Context Management Architecture in Pervasive Environments," *Personal and Ubiquitous Computing*, vol. 16, no. 6, 2012, pp. 741–755.

(b) N. Zhang and C. Todd, "Developing a Privacy Ontology for Privacy Control in Context-Aware Systems," Dept. of Electronic & Electrical Eng., Univ. College London, 2006; www.ee.ucl.ac.uk/lcs/previous/LCS2006/35.pdf.

(c) O. Sacco and A. Passant, "A Privacy Preference Ontology (PPO) for Linked Data," *Proc. Workshop Linked Data on the Web* (LDOW 11), 2011; iswc2011.semanticweb.org/fileadmin/iswc/Papers/Workshops/SPIM/spim2011_paper12.pdf.

(d) Y.-J. Hu and J.-J. Yang, "A Semantic Privacy-Preserving Model for Data Sharing and Integration," *Proc. Int'l Conf. Web Intelligence, Mining, and Semantics* (WIMS 11), 2011, article no. 9.

(e) M. Kost and J.C. Freytag, "Privacy Analysis Using Ontologies," *Proc. 2nd ACM Conf. Data and Application Security and Privacy* (CODASPY 12), 2012, pp. 205–216.

(f) M. Ahmed, A. Anjomshoaa, and A.M. Tjoa, "Context-Based Privacy Management of Personal Information Using Semantic Desktop: SemanticLIFE Case Study," *Proc. 10th Int'l Conf. Information Integration and Web-based Applications & Services* (iiWAS 08), 2008, pp. 214–221.

(g) I. Panagiotopoulos et al., "PROACT: An Ontology-Based Model of Privacy Policies in Ambient Intelligence Environments," *Proc. 14th Panhellenic Conf. Informatics* (PCI 10), 2010, pp. 124–129.

(h) M. Youssef, N.R. Adam, and V. Atluri, "Semantically Enhanced Enforcement of Mobile Consumer's Privacy Preferences," *Proc. ACM Symp. Applied Computing* (SAC 06), 2006, pp. 1172–1176.

(i) P. Bodorik, D. Jutla, and M.X. Wang, "Consistent Privacy Preferences (CPP): Model, Semantics, and Properties," *Proc. ACM Symp. Applied Computing* (SAC 08), 2008, pp. 2368–2375.

(j) D. Garcia et al., "Towards a Base Ontology for Privacy Protection in Service-Oriented Architecture," *Proc. IEEE Int'l Conf. Service-Oriented Computing and Applications* (SOCA 09), 2009; doi:10.1109/SOCA.2009.5410467.

(k) L. Kagal et al., "Authorization and Privacy for Semantic Web Services," *IEEE Intelligent Systems*, vol. 19, no. 4, 2004, pp. 50–56.

(l) M. Hecker, T.S. Dillon, and E. Chang, "Privacy Ontology Support for E-Commerce," *IEEE Internet Computing*, vol. 12, no. 2, 2008, pp. 54–61.

(m) L. Martimiano, M. Goncalves, and E. dos Santos Moreira, "An Ontology for Privacy Policy Management in Ubiquitous Environments," *Proc. IEEE Network Operations and Management Symp.* (NOMS 08), 2008, pp. 947–950.

collaborate with one another, get updates about each other's activities and interests, and exchange information efficiently while minimizing disruptions. Online services such as Google Wave gather context information about each person. Custom-defined XML schemas are used to model user privacy requirements, and the system users can decide which consumers are allowed to access their context information (such as location) at any given time. These policies protect this context information, which is only released to authorized personnel.

Ontology-based privacy-knowledge modeling

In an effort to develop a privacy ontology for context-aware systems, Ni Zhang and Chris Todd adopted P3P terminology and created corresponding classes and properties.[10] They defined a privacy rule class to represent privacy preferences set by users. Every privacy rule is expressed with two elements: data (data class) and conditions (condition class). The conditions class contains all conditions under which a user is willing to disclose data. As in the P3P specification, the conditions can be classified based on various individual preferences including data recipients, purposes of the data collection, duration that recipients will keep the data, a user's access privilege to his or her personal data once stored by recipients, and ways of handling disputes. With this approach, each data item (or collection) is attached to a rule that consists of a set of conditions.

Owen Sacco and Alexandre Passant proposed a lightweight vocabulary built on top of Web Access Control (WAC; www.w3.org/wiki/WebAccessControl) that enables users to create fine-grained privacy control for their data.[11] The access restrictions are put in place on an individual resource level (that is, the document level, not data items within the document level). WAC is a vocabulary that defines access control privileges in web documents.

Yuh-Jong Hu and Jiun-Jan Yang adopted a similar ontology where resources are protected at the resource level through conditions.[12] They model factors such as the allowed and not allowed uses of the data, time period for data retention, which data consumers are allowed and not allowed to access data, and obligations. One of the highlights in this work is that capturing is allowed and not allowed separately (such as users and purposes). This provides users with an additional opportunity when evaluating a given data request. As a result, global rules can be defined to allow a data request if a data owner has not explicitly defined certain factors. This might also combine with other factors. For example, one condition could be that if the data consumer is a research institute, data is allowed even if it has not defined the data retention period exactly.

Going a step further, Martin Kost and Johann Freytag identified 10 factors that need to be captured by privacy-preference modeling: purpose, consent, limited collection, limited use, disclosure, retention, accuracy and context preservation, security, openness, and compliance.[13] Compared with the other approaches, this work highlights the importance of capturing a broader range of information about a particular data collection and analysis task.

Mansoor Ahmed and his colleagues included a much broader set of privacy concepts in their knowledge model.[14] This work models knowledge that describes how data will be treated during the communication, transfer, storage, and processing of data. The PROACT ontology was designed for ambient environments and also captures detailed information about data-processing mechanisms, such as data collection, transfer, and storage.[15]

In their work, Mahmoud Youssef and his colleagues proposed a model to capture the privacy preferences of mobile consumers.[16] One of the significant features of this ontology is that it captures incentives. The authors analyzed different types of promotions and found five classes: monetary, coupon, time slack, extra items, and payment on installments. In relation to each incentive class, users can specify expected values as well.

Peter Bodorik and his colleagues also proposed a privacy preference model.[17] It lets users place restriction based on factors such as the purpose of data usage, data recipient, data retention, disputes, remedy, and access control (who has access to the data). Some of the purposes they list are admin, development, tailoring, pseudo-analysis, pseudo-decision, individual analysis, individual decision, contact, historical, and telemarketing. Similarly, some of the retention options are no retention, stated purpose, legal requirement, indefinitely, and business practices. Other approaches have presented similar retention details.[18,19]

Rei is a highly expressive policy language that lets users specify their privacy preferences.[20] An interesting concept presented in Rei is *rule priority*, which lets data owners define different combinations of conditions with different outcomes. Previous approaches only allowed one rule with multiple

> IN OPEN DATA MARKETS, IT IS IMPORTANT TO UNDERSTAND USERS' PRIVACY NEEDS PROACTIVELY AND PREDICT THEIR PREFERENCES AHEAD OF TIME.

conditions, where data consumers needed to meet all the conditions to access the data. Such expressions are important in data market negotiations, where users might want to define different sets of privacy preference conditions based on the type of data consumer. For example, a data owner might expect the data consumer to perform limited knowledge discovery if it is a commercial entity. However, the same data owner might grant unlimited knowledge discovery for not-for-profit medical research institutes.

In addition to the other factors modeled by the previously described privacy-knowledge model, Luciana Martimiano and her colleagues also included trust levels.[21] They established fixed sets of classes with static assignments such as close family, unknown person, known person, close friend, and coworker. As opposed to a Facebook-type categorization, where an individual can only be a friend or not a friend, their approach is much more aligned with how social interactions actually work. In a related work, Zahid Iqbal and his colleagues also highlighted the importance of modeling trust.[22]

LESSONS LEARNED

One of the important recent trends in this area is the increasing adoption of ontologies to model privacy knowledge.[10] An ontology defines a common vocabulary for researchers who need to share information in a domain, capturing the meaning between different concepts. It includes machine-interpretable definitions of basic concepts in the domain and the relations among them. Furthermore, ontologies promote the reuse of domain knowledge.

In some cases, researchers have used custom-defined XML schemes to model privacy.[7] For example, the P3P approach is driven by an XML schema. However, XML-like markup languages are typically used to structure data, but they cannot capture semantics and relationships. As we discussed earlier, one of the main weaknesses of the P3P approach was that it made it difficult to arrive at an agreed-upon vocabulary. Ontologies can address this issue because they allow for the modeling of relationships between similar terms. Therefore, the standardization of ontologies is not critical compared with markup-language-based modeling.

Another advantage of using ontologies to model privacy knowledge is that they allow privacy policies to be defined at both the data (instance) and class levels, which is more convenient for users. Instance-level rules should be given priority, and class-level rules could be used in the absence of instance-level rules.

Users' privacy preferences can change over time, so ideal systems should be able to adapt autonomously. In open data market scenarios, it is important to understand users' privacy needs proactively and predict their preferences ahead of time so the data owners do not have to deal with privacy configurations.

One of the common weaknesses in current approaches is the lack of support to capture and model information about data-management techniques. For example, what techniques are being used to store data (such as encryption techniques), and how will the data be routed (see torproject.org)? These are important factors for data owners that might directly affect whether they choose to share data with a particular data consumer. Even though some approaches have highlighted the importance of modeling data-management-related information, researchers have yet to introduce proper technique-level vocabulary and concepts.

In general, most approaches outline a privacy policy in some form that includes rules, preferences, conditions, and so on. Policies typically define who can access a certain resource and under which conditions, how data should be provided to data consumers, how the provided information will be used, and so on. If done at the data-item level, such privacy preference configurations could be exhaustive. However, the use of ontologies makes this simpler by supporting class-level policy definitions.

Based on a number of privacy regulations, a study by Diego Garcia and his colleagues outlined a set of privacy requirements that should be considered when developing a privacy-knowledge model:[18]

- A description of the data collected and how it is used by consumers should be available to data owners.
- Data owners should be able to agree with the collection of their data before it happens.
- The techniques used to collect a data item should be identified.
- The collector (data consumer) of a data item should be identified.
- The purposes for which a data item is collected should be identified.
- The entities (a third party, for example) to which a data item is disclosed by its collector should be identified.
- The data items to be collected should be identified.
- The retention time of a data item should be identified.

- Data consumers should indicate if data owners are allowed to complete, correct, and update their retained data.
- Data consumers should indicate if data owners can request records on how their data have been used, in formats understandable by data owners and with known delays and charges.
- Data consumers should indicate if data owners are able to request copies of data on them, in formats understandable by data owners and with known delays and charges.

We expanded the researchers' requirements list based on the lessons learned by evaluating the approaches included in Table 2:

- Data consumers should clearly inform data owners regarding what kind of knowledge is expected to be discovered using their data.
- Data owners should know the risks—their impact level of sharing (trading)—regarding a particular type of data before sharing (trading) occurs.
- Data owners and consumers should come to an agreement regarding the reward that the data owners might receive as a return for taking the risks of sharing data.
- Reward types associated with sharing data need to be identified clearly before any data sharing occurs.
- Data owners should be able to apply data-quality-reduction techniques before data is sent to the data consumers to reduce privacy risks.
- Data owners and consumers should agree on which data-quality-reduction techniques will be used.

Thus, our expanded list also recommends ways to support the needs of the IoT and open data markets.

FUTURE RESEARCH DIRECTIONS, CHALLENGES, AND OPPORTUNITIES

Here, we briefly discuss some of the major research challenges that need to be addressed in the future, with a particular emphasis on the needs of the IoT[1] and open data markets.[5]

Privacy-preference modeling and user profiling

Privacy-preference profiling is the task of modeling user preferences in a common structure. Many factors could impact a user's privacy preferences, especially in the IoT domain and open data market scenarios. One major challenge is to identify all the factors that could impact users' decisions when they think about their privacy expectations. For example, when participating in open data markets, some users (data owners) might consider the data consumer's reputation to be the most important factor when deciding whether to trade data. For other users, the purpose of the data collection could be the most significant factor. At the same time, some factors could be completely meaningless to some users, depending on their level of technical knowledge. For example, the type of encryption supported by a given data consumer might not impact nontechnical users' decisions because they have no way to evaluate and understand the value of encryption.

The advantage of modeling each user's privacy knowledge is that it allows both humans and machines to share a common vocabulary. For example, in highly dynamic environments such as the IoT, the automated configuration of machine-interpretable privacy preferences could significantly reduce users' workload and privacy concerns. Furthermore, a common understanding will help different software programs use the privacy-knowledge model to provide different types of value-added services, such as proactive privacy-preference configuration, learning user behavior over time, and predicting users' privacy expectations. Additionally, a common privacy-preference knowledge model would be helpful in conducting data-trading negotiations in open data market environments.[5]

An ideal privacy-knowledge model should be able to capture any piece of information that could potentially impact privacy. Such models should be able to capture users' priorities, wherein each user can value various factors differently.

User privacy-preference acquisition

Once we have a privacy-preference knowledge model, which can be considered a template, the next challenge is to acquire user (data owners) privacy preferences with minimal human intervention. Asking for too much information about preferences from data owners might overload them, whereas a lack of information could lead to violations of their privacy expectations.

Recommender systems could help address this issue. For example, we could build a basic template for each user by analyzing and studying similar users based on demographics. Using this approach could allow us to identify users' personalities with a limited question-and-answer mechanism. Then, recommender systems could predict some parts of the privacy preferences and question users again to fill the remaining essential privacy-preference parameters.

From the data consumers' perspective, one of the main challenges is the costs associated with and scalability of data acquisition. Data-acquisition negotiations must be done individually with each data owner. Even though a single data-trading transaction might not consume substantial amounts of computational resources, the costs will scale exponentially for large numbers of such transactions. Therefore, data-acquisition negotiation algorithms must be efficient and scalable.

Based on this survey of privacy-knowledge modeling techniques, concepts, and challenges, we can see that the IoT domain demands a more comprehensive privacy-knowledge modeling approach. We recommend the development of a comprehensive knowledge model that is capable of capturing user privacy knowledge. Specifically, we argue that ontologies represent the most appropriate method of modeling privacy knowledge because of their ability to model relationships between concepts and support of automated reasoning. Furthermore, a key area of future research will be in developing techniques that can acquire privacy preferences autonomously, with limited intervention from users, to avoid overloading them.

ACKNOWLEDGMENTS

Charith Perera's work is funded by European Research Council advanced grant 291652 (ASAP). Albert Zomaya's work is supported by an Australian Research Council Discovery grant (DP130104591).

REFERENCES

1. C. Perera et al., "Context Aware Computing for The Internet of Things: A Survey," *IEEE Comm. Surveys Tutorials*, vol. 16, no. 1, 2013, pp. 414–454.
2. W. Ren, "QoS-aware and Compromise Resilient Key Management Scheme for Heterogeneous Wireless Internet of Things," *Int'l J. Network Management*, vol. 21, no. 4, 2011, pp. 284–299.
3. W. Ren, "uLeepp: An Ultra-Lightweight Energy-Efficient and Privacy-Protected Scheme for Pervasive and Mobile WBSN-Cloud Communications," *Ad Hoc & Sensor Wireless Networks*, vol. 27, nos. 3–4, 2015, pp. 173–195.
4. C. Perera, C. Liu, and S. Jayawardena, "The Emerging Internet of Things Marketplace from an Industrial Perspective: A Survey," *IEEE Trans. Emerging Topics in Computing*, vol. 3, no. 4, 2015, pp. 585–598.
5. C. Perera, R. Ranjan, and L. Wang, "End-to-End Privacy for Open Big Data Markets," *IEEE Cloud Computing*, vol. 2, no. 4, 2015, pp. 44–53.
6. A.F. Westin, *Privacy and Freedom*, Bodley Head, 1967.
7. L.F. Cranor et al., "2006 Privacy Policy Trends Report," CyLab, Carnegie Mellon Univ., 2007.
8. Electronic Privacy Information Center, "Pretty Poor Privacy: An Assessment of P3P and Internet Privacy," 2000; epic.org/reports/prettypoorprivacy.html.
9. Z. Jaroucheh, X. Liu, and S. Smith, "An Approach to Domain-Based Scalable Context Management Architecture in Pervasive Environments," *Personal and Ubiquitous Computing*, vol. 16, no. 6, 2012, pp. 741–755.
10. N. Zhang and C. Todd, "Developing a Privacy Ontology for Privacy Control in Context-Aware Systems," Dept. of Electronic & Electrical Eng., Univ. College London, 2006; www.ee.ucl.ac.uk/lcs/previous/LCS2006/35.pdf.
11. O. Sacco and A. Passant, "A Privacy Preference Ontology (PPO) for Linked Data," *Proc. Workshop Linked Data on the Web* (LDOW 11), 2011; iswc2011.semanticweb.org/fileadmin/iswc/Papers/Workshops/SPIM/spim2011_paper12.pdf.
12. Y.-J. Hu and J.-J. Yang, "A Semantic Privacy-Preserving Model for Data Sharing and Integration," *Proc. Int'l Conf. Web Intelligence, Mining, and Semantics* (WIMS 11), 2011, article no. 9.
13. M. Kost and J.C. Freytag, "Privacy Analysis Using Ontologies," *Proc. 2nd ACM Conf. Data and Application Security and Privacy* (CODASPY 12), 2012, pp. 205–216.
14. 1M. Ahmed, A. Anjomshoaa, and A.M. Tjoa, "Context-Based Privacy Management of Personal Information Using Semantic Desktop: Semantic-LIFE Case Study," *Proc. 10th Int'l Conf. Information Integration and Web-based Applications & Services* (iiWAS 08), 2008, pp. 214–221.
15. I. Panagiotopoulos et al., "PROACT: An Ontology-Based Model of Privacy Policies in Ambient Intelligence Environments," *Proc. 14th Panhellenic Conf. Informatics* (PCI 10), 2010, pp. 124–129.

ABOUT THE AUTHORS

CHARITH PERERA is a research associate at the Open University, where he is working on the Adaptive Security and Privacy (ASAP) research program. His research interests include the Internet of Things (IoT), sensing as a service, privacy, middleware platforms, and sensing infrastructures. Perera received a PhD in computer science from the Australian National University. He is a member of IEEE and ACM. Contact him at charith.perera@ieee.org.

CHANG LIU is a postdoctoral research associate in the Data61 Group at CSIRO. His research interests include cloud computing, big data, data security, data privacy, and applied cryptography. Liu received a PhD in computer science from the University of Technology Sydney. Contact him at chang.liu@csiro.au.

RAJIV RANJAN is an associate professor in the School of Computer Science at the China University of Geosciences and in the School of Computing Science at Newcastle University. His research interests include cloud computing, the IoT, big data, distributed systems, and peer-to-peer networks. Ranjan received a PhD in computer science from the University of Melbourne. He serves on the editorial boards of *IEEE Transactions on Computers*, *IEEE Transactions on Cloud Computing*, *IEEE Cloud Computing*, and *Future Generation Computer Systems*. He is a member of IEEE. Contact him at raj.ranjan@ncl.ac.uk.

LIZHE WANG is a ChuTian Chair Professor in the School of Computer Science at the China University of Geosciences and a professor in the Institute of Remote Sensing and Digital Earth at the Chinese Academy of Sciences. His research interests include high-performance computing, e-science, and remote sensing image processing. Wang received a PhD in engineering from the University of Karlsruhe. He is an associate editor of *IEEE Transactions on Computers*, *IEEE Transactions on Cloud Computing*, and *IEEE Transactions on Sustainable Computing*. Wang is a Fellow of the Institute of Engineering and Technology and the British Computer Society. Contact him at lizhe.wang@gmail.com.

ALBERT Y. ZOMAYA is a chair professor and director of the Center for Distributed and High Performance Computing at Sydney University. He is the editor in chief of *IEEE Transactions on Sustainable Computing* and serves as an associate editor for 22 leading journals. He is the recipient of the IEEE TCPP Outstanding Service Award, the IEEE TCSC Medal for Excellence in Scalable Computing, and the IEEE Computer Society Technical Achievement Award. Zomaya is a Fellow of the American Association for the Advancement of Science, IEEE, and the Institute of Engineering and Technology. Contact him at albert.zomaya@sydney.edu.au.

16. M. Youssef, N.R. Adam, and V. Atluri, "Semantically Enhanced Enforcement of Mobile Consumer's Privacy Preferences," *Proc. ACM Symp. Applied Computing* (SAC 06), 2006, pp. 1172–1176.
17. P. Bodorik, D. Jutla, and M.X. Wang, "Consistent Privacy Preferences (CPP): Model, Semantics, and Properties," *Proc. ACM Symp. Applied Computing* (SAC 08), 2008, pp. 2368–2375.
18. D. Garcia et al., "Towards a Base Ontology for Privacy Protection in Service-Oriented Architecture," *Proc. IEEE Int'l Conf. Service-Oriented Computing and Applications* (SOCA 09), 2009; doi:10.1109/SOCA.2009.5410467.
19. M. Hecker, T.S. Dillon, and E. Chang, "Privacy Ontology Support for E-Commerce," *IEEE Internet Computing*, vol. 12, no. 2, 2008, pp. 54–61.
20. L. Kagal et al., "Authorization and Privacy for Semantic Web Services," *IEEE Intelligent Systems*, vol. 19, no. 4, 2004, pp. 50–56.
21. L. Martimiano, M. Goncalves, and E. dos Santos Moreira, "An Ontology for Privacy Policy Management in Ubiquitous Environments," *Proc. IEEE Network Operations and Management Symp.* (NOMS 08), 2008, pp. 947–950.
22. Z. Iqbal et al., "Toward User-Centric Privacy-Aware User Profile Ontology for Future Services," *Proc. 3rd Int'l Conf. Comm. Theory, Reliability, and Quality of Service* (CTRQ 10), 2010, pp. 249–254.