
A User Interface for Encoding Space Usage Rules Expressed in Natural Language

Pavel Samsonov

Expertise Centre for Digital Media
Hasselt University - tUL - iMinds
pavel.samsonov@uhasselt.be

Johannes Schöning

Expertise Centre for Digital Media
Hasselt University - tUL - iMinds
Johannes.schoening@uhasselt.be

Brent Hecht

Computer Science and Engineering
University of Minnesota
bhecht@cs.umn.edu

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.
Copyright is held by the owner/author(s).
CHI'15 Extended Abstracts, Apr 18-23, 2015, Seoul, Republic of Korea
ACM 978-1-4503-3146-3/15/04.
<http://dx.doi.org/10.1145/2702613.2732737>

Abstract

Our interactions with the spaces around us are frequently defined by *space usage rules* (SURs) like “no smoking”, “no dogs allowed”, and “stay on the trail”. These rules are important public health tools and help protect the environment, among other applications. However, despite their importance, no large-scale database of SURs currently exists. This prevents online/mobile maps from presenting these rules to users, as their traditional paper counterparts have been shown to do regularly. The lack of a SUR database also prevents developers from building novel SUR-based applications, e.g. mobile apps that inform users where they can smoke or walk their dog. In this paper, we present an in-development user interface to support the semi-automatic encoding of SURs expressed in natural language in documents such as laws, city ordinances, park rules, and institution FAQ pages. Our system will allow an untrained user to easily enter complex rules into a spatial data format compatible with OpenStreetMap.

Author Keywords

Space usage rules; context-aware applications; interactive maps.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous. User interfaces; Graphical user interfaces (GUI); Input devices and strategies; Interaction styles.

Introduction

Recent work (e.g. [9]) strongly argues for the importance of space usage rules (SURs) in location-aware technologies. SURs, which are broadly defined to be rules and regulations that are spatially-constrained in nature, range from everyday restrictions like ‘no smoking’ and ‘no dogs allowed’ to less common restrictions like ‘stay on the trail’ and ‘no campfires’. Informing users of SURs in a specific area would enable an entirely new class of context-aware intelligent interfaces. For instance, imagine a mobile app [8] that informs users if it is legal to light a cigarette in the current location, and, similarly, a mobile app that tells hunters if it is okay to start a campfire. One can also easily imagine straightforward algorithms that provide routing instructions to help dog owners avoid “no dogs allowed” areas when walking their dogs and those that generate vacation recommendations for specific areas that allow activities of interest (e.g. climbing, fishing, swimming).

In a recent study [7], we showed that the main obstacle to the inclusion of SURs in online and mobile maps and the development of the aforementioned applications is the lack of available SUR data. The only open source SUR datasets that exist are those that can

be derived from tag information in OpenStreetMap¹ (OSM). While the vector geographical information in OSM has extensive coverage and is of high quality in many places (e.g. [2] and [5]), tag data in OSM such as “no smoking” and “no dogs” is very sparse. For instance, in our recent work [7], we reported that only 47 water bodies in the entire world in OSM have been tagged with “no fishing”. Applying these tags requires a lot of manual effort and they are rarely used.

To address this sparsity of SUR data and to help enable the location-aware technologies mentioned above (and many others), we are developing a user interface for encoding in machine-readable format the thousands of SURs expressed in natural language in documents such as legal statutes, park rules, and institution FAQ pages. Natural language is hard to interpret by a computer and ambiguities may occur. Therefore, rules are processed by users and stored as script programming code in the Lua² language. Users do not have to learn the Lua syntax or do not require any programming skills, as they can use our wizard-styled “if-this-then-that” user interface that is in the early stages of development. This application creates Lua code that is used to encode SUR data. The interface also allows users to specify additional options, such as the dates or time of the day when an activity is restricted or forbidden, or additional specifications, for example, when dog walking is allowed only if the dog is leashed and under control of its owner at all times. Afterwards users may display their SURs encoded in LUA on a map. The proposed system has the potential to help, for instance, crowdworkers who are mapping an OSM-based tag

¹ www.openstreetmap.org

² www.lua.org

database. Imagine a rule from the California state parks website, which tells us that dogs are prohibited on all trails and beaches in California state parks. Using the “if this than that” interface, converting this rule into Lua becomes a trivial task for users, while manually adding OSM tags to all the beaches and trails in more than 50 state parks in California requires a lot of work. Moreover, if the rule changes, editing the rule in the interface will apply the new version, while manually applying tags in OSM requires the user to do the same work again.

Related Work

In the OpenStreetMap tag database, a variety of SURs are present, such as “no access”, “no horse riding”, “no dog walking”, and many others. OSM allows users to edit SURs on the map by applying tags to elements, such as buildings, areas, roads and points of interest. However, as was noted above, the quality and quantity of SUR tag data is rather low. Also, there are many possible SURs that are not present in the OSM system, such as “no drone flying” or “no swimming”. Complex rules that refer to time of day, date or special conditions cannot be tagged in OSM. Our system enriches the OSM SUR dataset with data extracted from rules and can visualize any desired SURs of any complexity on the map.

The only other research on the subject of mapping SURs is a computer vision algorithm we presented in the past [7]. The algorithm can detect a variety of restriction signs in digital pictures, such as “no smoking”, “no fishing” etc. This algorithm was used to process pictures on the photo sharing website Flickr³

³ www.flickr.com

and found hundreds of space usage rules that were not already tagged in OSM. However, many SURs cannot be or are not posted on signs. Consider, for instance, this rule from the Alaska State Park website⁴: “Discharge of firearms within 1/2 mile of any developed park facility is prohibited”. For a hunter, this rule can be difficult to follow, as it is sometimes hard to find out if there are developed park facilities in 1/2 mile around (e.g. in a forest), and some people might not even be familiar with this rule. Placing signs that say “no firearm discharging behind this point” everywhere around each park facility is unrealistic. Using the *rule-based tagging approach* of our LUA-based system solves this problem. Using our system, a person (e.g. a crowdworker) can encode this natural language rule in LUA and this information can then be used to support location-aware applications that inform hunters of where it is legal or illegal to hunt.

Rule-based Approach Pipeline

The pipeline used by our system can be seen in Figure 1. First, a rule is extracted from a webpage on the Internet. The rule is then analyzed to find out which areas in the world are affected by those rules (regions of interest, or ROI).

⁴ <http://dnr.alaska.gov/parks/faq.htm>

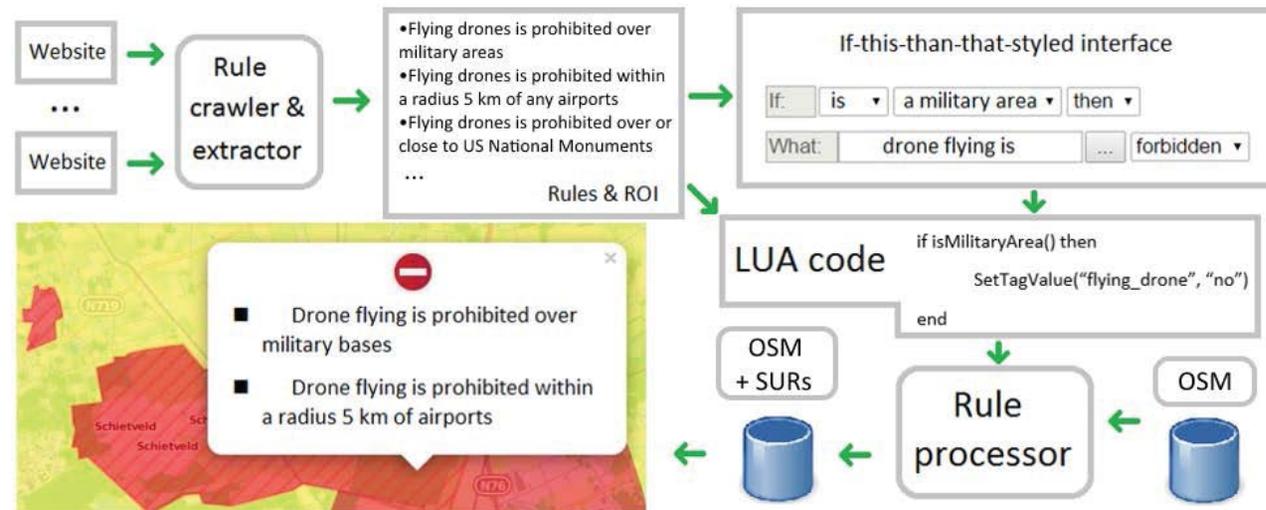


Figure 1. Rule editing process and application example built using our proposed system. The Lua functionality is completed, but the “if-this-than-that”-style interface is still in development.

In most cases, the area name is unique and exists in OSM, and the user may just refer to it by name, but in some rare cases the user has to pick the affected polygon on the map manually. Rule extraction and location (i.e. the part of the pipeline described in this paragraph) is currently done manually, but in the future, this process can be automated using lightweight natural processing and *geoparsing* (e.g. [1], [3]).

Next, rules are converted from natural language to strict machine-understandable Lua scripting language. Users can directly program in Lua or use the “if-this-then-that” interface, which is currently in development. If the interface is used, the user selects if the rule applies to a whole territory or to a sub-area, and chooses, which

activities are forbidden or restricted on the specified territory. For restricted activities, users are able to choose the time of the day or dates when the activity is forbidden and additional options, such as “only leashed dogs”, “no LPG vehicles”, and many others. Users can construct AND/OR combinations in any form from simpler rules as well. This is how the rule-based tag dataset is filled with SUR data. Next, the rule processor applies rules in Lua to specified regions of interest using geographical information from OSM and creates the final database with all the geographical information and SURs. This step is fully automatic. In Figure 2, an example of an area around Central Park in New York is shown. Red zones show where it is forbidden to consume alcohol. As alcohol consumption is forbidden in all the park as stated

Rules for Figures 2, 3:

“Per the New York City Department of Parks & Recreation, which establishes the rules and regulations for all City parks, alcohol is prohibited in Central Park”.

“Dogs are welcome in Central Park. They may be off-leash between the hours of 9:00 pm and 9:00 am in most sections of the Park, but must be under the control of their owner at all times. Dogs are prohibited in all ballfields and recreational courts, playgrounds, water bodies, streams, ornamental fountains, elm islands at the Mall and the Great Lawn Oval, Lilac Walk, and Reservoir running track. Dogs must be leashed at all times in the following locations: Arthur Ross Pinetum, The Bridle Path, Cedar Hill, Children's Glade (Great Hill area), Conservatory Garden, East Green, East Meadow Oval, Kerbs Boathouse Plaza, The North Woods and The Ravine, The Ramble, Shakespeare Garden, Strawberry Fields and Turtle Pond Lawn.”



Figure 2. Map of Central Park, NY, showing areas in the park where alcohol consumption is forbidden.



Figure 3. Map of Central Park, NY, showing areas in the park where dog walking is restricted (blue) or forbidden (red).

on Central Park’s website, all its territory is painted in red. In Figure 3, red zones show where dog-walking is completely forbidden and blue zones show where it is restricted by time of the day or where dogs should be kept on a leash. Some sub-areas in OSM do not exist yet, for example Great Lawn Oval, and therefore this

area is not marked in the map. However, as the OSM database becomes more and more precise and consistent with the help of people contributing information to it [6], in the future this area will be tagged by rules as well. Also, most of the park area is normally painted in blue as dogs should be kept leashed

all night, but only a blue edge was left for better clarity in the example. In the red areas no dogs are allowed.

Conclusion & Discussion

In this paper, we explained how data about space usage rules, such as “no smoking” or “do not fly your drone here”, can enable a new class of location-aware applications and we described a human-in-the-loop system for gathering this data from natural language rules. This system is only semi-automatic, as it requires manual work to extract the rules from the Web and convert them to Lua code. However, we discussed how using the system is more efficient than tagging each building, road or park facility manually.

In the future, we want to set up a server to make the SUR data accessible for everyone. The server application will download recent geographical data from the OSM server and use it to generate custom feature-based tag data. To fill the database with rules, we plan to use Mechanical Turks. We will make a rule extractor that analyzes websites on the Internet to extract potential rules. Each rule will be given to several crowdworkers to convert into Lua code using the “if-this-than-that” interface, and if the crowdworkers’ encodings of the rule agree, the SUR will be entered into the database.

Finally, another feature that we want to implement for the server application, is applying rules to the map on the fly. Imagine a legislator developing new laws to prohibit flying drones or swimming too far from the coast. By using the “if-this-than-that” interface and changing the radius of prohibited flights or the maximum distance from the coast, the legislator will immediately be able to see on the map how various versions of the law would affect her or his districts.

References

- [1] Gelernter, J., and Mushegian, N. Geo-parsing Messages from Microtext. *Transactions in GIS 15.6 (2011): 753-773.*
- [2] Haklay, M. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and planning. B, Planning & design 37.4 (2010): 682.*
- [3] Hoffart, J., Yosef, M., Bordino, I., Fürstenau, H., Pinkal, M., Spaniol, M., Taneva, B., Thater, S. and Weikum, G. Robust disambiguation of named entities in text. In *Proc. of EMNLP (2011).*
- [4] Ierusalimschy, R., Henrique de Figueiredo, L., Celes, W. The Evolution of Lua. In *Proc. of the 3rd ACM SIGPLAN Conference on History of Programming Languages, HOPL III, pp. 2-1-2-26. ACM, New York (2007).*
- [5] Mooney, P., Corcoran, P., Wistanley, A.C. Towards Quality Metrics for OpenStreetMap. In *Proc. of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, ACM (2010) 514–517.*
- [6] Mordechai, H. and Weber, P. OpenStreetMap: User-Generated Street Maps. *Pervasive Computing, IEEE (2008). Volume: 7, Issue: 4.*
- [7] Samsonov, P., Tang, X., Schöning, J., Kuhn, W. and Brent, H. You Can’t Smoke Here: Towards Support for Space Usage Rules in Location-aware Technologies. In *Proc. of CHI. (2015).*
- [8] Schall, G., Schöning, J., Paelke, V., & Gartner, G. A survey on augmented maps and environments: approaches, interactions and applications. *Advances in Web-based GIS, Mapping Services and Applications (2011).*
- [9] Schöning, J., Hecht, B. and Kuhn, W. Informing Online and Mobile Map Design with the Collective Wisdom of Cartographers. In *Proc. of DIS '14 (2014).*