Mechanismen zur Beschaffung korrekter Daten

Boi Faltings und Goran Radanovic
EPFL-LIA
boi.faltings@epfl.ch, goran.radanovic@epfl.ch

September 12, 2016

Abstract

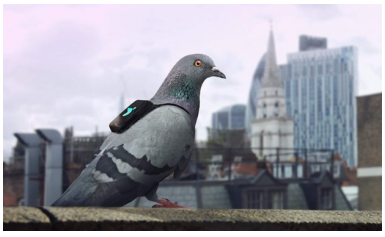
Smart Cities brauchen als Grundlage für Entscheidungen grosse Mengen von Daten, die häufig nicht direkt, sondern über andere Agenten erhoben werden. Es ist daher wichtig, die Qualität durch geeignete Anreize oder Kontrolle sicherzustellen. Bei Ausnutzung der Spieltheorie ist dies erstaunlicherweise selbst dann möglich, wenn die Korrektheit der Daten gar nicht kontrolliert werden kann.

Wir zeigen neuartige Mechanismen zur Qualitätssicherung für zwei Arten von Agenten. Für rational handelnde Agenten berechnen wir die Belohnung für die Daten so, dass nur korrekte Daten einen positiven Betrag erwarten können. Für fehlerhafte oder absichlich falsche Daten zeigen wir einen neuen Mechanismus, der den negativen Einfluss auf das erzeugte Modell begrenzt.

1 Einführung

Die Grundlage von "Smart Cities" besteht darin, dass aufgrund von kontinuierlich erhobenen Daten die Funktionen und Regeln der städtischen Dienste laufend an die momentane Situation angepasst werden. So könnte z.B. der Zustrom von Verkehr in die einzelnen Regionen der Stadt in Echtzeit von der momentan gemessenen Luftverschmutzung und Lärmbelastung abhängig gemacht werden. Eine wesentliche Vorraussetzung für solche Automatismen ist die kontinuierliche Erhebung von Daten: in diesem Beispiel muss die Luftverschmutzung kontinuierlich und räumlich verteilt gemessen werden. Neuere Forschung hat nämlich gezeigt, dass die Luftverschmutzung räumlich sehr stark variiert ([1, 2]), so dass eine zentrale Messung, wie sie heute praktiziert wird, wahrscheinlich kein korrektes Bild wiedergibt.

Nun ist es für staatliche Stellen sehr aufwendig, selber überall solche Sensoren aufzustellen und auch zu warten. Wesentliche einfacher wäre es,



Pigeon patrol

Plume Sensor

Air Quality Egg

Laser Egg

Figure 1: Beispiele von kostengünstigen Sensoren für Luftverschmutzung.

wenn man diese Erhebung den Bürgern selber überlassen könnte, so wie z.B. WAZE ([3]) Daten über die Verkehrslagen direkt von den Verkehrsteilnehmern bezieht. Auf diese Art wäre auch eine räumliche Verteilung erreichbar, die in etwa derjenigen der Bevölkerung entspricht. Des weiteren könnten Bürger, die dem Staat mistrauen, selber messen und diese Messwerte auch in den öffentlichen Prozess einbringen. Damit kann eine staatliche Beschönigung der Luftverschmutzungswerte, wie sie weitgehend vermutet wird ([4, 5]), verhindert werden. Dieser Gesichtspunkt einer demokratischen Kontrolle könnte für die Akzeptanz von Smart Cities in der Bevölkerung wichtig sein.

Waren hinreichend genaue Messungen von Luftverschmutzung aufgrund der geringen Konzentrationen bisher nur mit sehr aufwendigen Geräten möglich, so kann man heute mit den Techniken des Internets der Dinge schon zu wesenlich günstigeren Kosten recht genaue Messungen von NOx, CO und Feinstaub erhalten. Forschungsprojekte wie das Opensense-Projekt, haben gezeigt, dass auch mit low-cost Sensoren eine recht gute Messqualität erzielt werden kann ([6]), die mit geeigneten Modellierungstechniken noch weiter erhöht werden kann ([7]).

Abbildung 1 zeigt einige Beispiele von low-cost Sensoren, die heute auf dem Markt sind. Im Jahre 2012 wurde das Air Quality Egg ([8]) als "Best of Kickstarter" Projekt ausgezeichnet, wobei allerdings die Messqualität miserabel war. Neuerdings gibt es ähnliche Sensoren auch in sehr kleiner, mobiler Ausführung, die sogar von Tauben transportiert werden (Pigeon Air Patrol [9]). Eine erheblich bessere Qualität zu geringen Kosten kann mit Lasern erzielt werden, wie sie kürzlich mit dem Laser Egg ([10]) auf den Markt gebracht wurden.

Somit kann in Kürze mit einer breiten Verfügbarkeit von - bei sachgemässen Einsatz - relative hochwertigen Messinstrumenten für Luftqualität gerechnet werden. In anderen Bereichen, wie z.B. Messung von Lärm, ist schon heute eine einfache Messung durch die Bürger selbst möglich. Nun stellt sich die Frage, wie diese Messdaten in sinnvoller Weise gesammelt und benutzt werden können, was unter dem Namen Crowdsensing bekannt ist. Wir werden in diesem Beitrag anhand des Beispiels der Luftverschmutzung zeigen, wie so etwas konkret implementiert werden könnte.

2 Probleme bei der Einführung von Crowdsensing

Das Betreiben eines Sensors für Luftqualität wird immer mit einem gewissen Aufwand verbunden sein. Neben den Anschaffungskosten, die auch in Zukunft wohl über 100 Euro liegen werden, muss ein solcher Sensor auch richtig aufgestellt, gereinigt und gelegentlich kalibriert werden. Es wir daher nötig sein, dass man diejenigen, die Sensordaten liefern, dafür auch in irgendeiner Weise belohnt. Dieses könnte dann Anschaffung und Betrieb eines Sensors zu einer gewinnträchtigten Aktivität machen, ähnlich der Einspeisung von Solarstrom.

Ein wesentlicher Unterschied zum Solarstrom besteht allerdings in der Qualitätskontrolle: während diese beim Strom sehr einfach ist, gibt es bei Daten keine einfache Kontrollmöglichkeit. Wenn jede Einspeisung von Daten bezahlt wird, dann könnte jeder einfach Zufallszahlen einspeisen und sich so die Kosten für die Messungen sparen! Daher ist das erste Problem dasjenige, ein Kriterium zu finden, was zumindest im Mittel nur echte Messwerte belohnt und alle anderen Strategien mit keinem oder negativem Erlös versieht. Damit wird dann auch dafür gesorgt, dass sich nur ernsthafte Agenten beteiligen (self-selection), was wohl von allen Massnahmen den grössten Einfluss auf die Datenqualität hat.

Wenn die Datenerfassung dazu erfolgt, automatische Massnahmen auszulösen, wie es ja bei Smart Cities der Fall ist, dann besteht des weiteren ein Anreiz,

die Daten zwecks einer Beeinflussung dieser Massnahmen zu verfälschen. Zum Beispiel kann man seine Luftverschmutzung dadurch verschleiern, dass man geschönte Messdaten liefert. Ein solches Verhalten kann auch unbeabsichtigt auftreten, wenn z.B. der Sensor defekt ist. In beiden Fällen reagiert der Sensor nicht auf monetäre Anreize, und man muss daher die gelieferten Daten durch einen anderen Prozess erkennen. Dieses ist dadurch möglich, dass man jedem Datenlieferant einen zahlenmässig quantifizierten Leumund (Reputation) zuordnet, der von der Qualität der bisher gelieferten Daten abhängt. Daten von Lieferanten mit schlechtem Leumund können dann bei der Entscheidungsfindung nachranging behandelt werden.

Darüberhinaus können natürlich auch statistische Methoden zum Filtern von Ausreissern (Outliers) und Kalibrierung mit den obengenannten Methoden kombiniert werden, um eine bestmöglichste Datenqualität zu gewährleisten. Das Ziel ist es, die Qualität des Gesamtsystems - Sensorik, Datensammlung, Modellierung und Entscheidungsfindung - zu optimieren. Dabei sind Anreize (Incentives) ein sehr attraktives Mittel, da sie die Möglichkeit bieten, falsche Daten vollständig zu eliminieren: wenn kein Anreiz besteht, solche Daten zu liefern, dann wird sich auch niemand die Mühe machen. Statistische Filter können falsche Daten nie vollständig eliminieren, sondern nur ihren Einfluss reduzieren. Dagegen reicht es für Anreize aus, dass die Agenten *glauben*, dass nur echte Daten belohnt werden, damit *alle* die bestmöglichsten Daten liefern. Solange die unvermeidbaren Fehler bei der Bewertung nicht die Überhand gewinnen, werden sie auf diesen Glauben keinen Einfluss haben. Man kann also eine Bewertung, die nur im Durchschnitt korrekt ist, zu einem hundertprozentig korrektem Datenfilter machen, sofern die Agenten darauf rational reagieren.

Es gab daher schon einige Arbeiten, die diesen Ansatz verfolgt haben, z.B. [11, 12]. In diesem Beitrag werden wir neuere Mechanismen vorstellen, die diese Resultate weiter verbessern.

3 Verifizierbare Messungen

Wir betrachten als erstes den Fall, wo die korrekten Daten zu einem späteren Zeitpunkt bekannt werden, und es sich bei den erhobenen Daten um *Vorhersagen* handelt. Dieses tritt in der Praxis recht häufig auf, wie folgende Beispiele zeigen. Wir können recht einfach durch Messung des Pegels in einem offenen Behälter den gesamten Regenfall innerhalb eines Tages messen und dieses dazu benutzen, die Korrektheit einer Reihe von kurzfristigen Messungen zu verifizieren. Nach dem gleichen Prinzip können wir Luftver-

schmutzung durch Partikel über längere Zeit kostengünstig durch die Akkumulation auf einem Filter verifizieren. Wenn wir Pilzbefall auf Nutzpflanzen erheben, können wir die Daten später durch den Erfolg der Ernte verifizieren.

Selbst wenn diese Grundwahrheit (ground truth) erst später bekannt wird, kann man sie sowohl für die Berechnung einer Belohnung (die dann halt erst später erfolgt) wie auch für die Zuweisung eines Leumundes (Reputation) benutzen. Wir werden nun beide Möglichkeiten betrachten.

3.1 Mechanismen für Anreize

Da es darum geht, Anreize für Wohlverhalten des Sensorbetreibers zu schaffen, sollten wir zunächst das System von seiner Seite her ansehen. Der Betreiber kann zwischen verschiedenen Strategien wählen, wie zum Beispiel verschiedene *heuristische* Strategien: immer den gleichen Wert melden, Zufallszahlen melden, oder den Messwert von einem anderen Sensor kopieren.

Im Gegensatz dazu wünschen wir uns aber eine *kooperative* Strategie, die darin besteht, eine sorgfältige Messung vornehmen und das genaue Ergebnis melden.

Das Ziel des Mechanismums ist es, dass der Betreiber eine kooperative Strategie wählt. Er sucht dazu, durch die versprochene Belohnung den Betreiber zu einem solchen Verhalten zu bewegen. Wir wollen insbesondere folgendes Verhalten hervorrufen:

1. der Betreiber hat für die möglichen Messwert eine Priori-Wahrscheinlichkeit $q(x)$, wobei $x_0 = \arg\max q(x)$ der wahrscheinlichste Wert ist.

2. der Betreiber macht eine Messung und erhält daraus eine Posteriori-Wahrscheinlichkeit $p(x)$, wobei $x_1 = \arg\max p(x)$ der wahrscheinlichste Wert ist.

3. der Betreiber meldet den Wert x_1.

Ganz wichtig ist nun die Annahme, dass sich der Betreiber *rational* verhält: er wählt diejenige Strategie, die ihm den höchsten Gewinn verspricht. Mit "Gewinn" meinen wir hier die Differenz Belohnung minus Aufwand. Das Wort "verspricht" ist entscheidend: es zählt, was der Betreiber glaubt, zu erhalten, und nicht, was er effektiv erhält. Dieses erleichtert unsere Arbeit ungemein, denn es ist hinreichend, dass wir *im Mittel* eine ausreichende Belohnung erzielen, um *alle* rationalen Betreiber bei *jeder* Messung zu korrektem Verhalten zu bewegen.

Im Folgenden nehmen wir an, dass jeder Sensor von einem rationalen *Agenten* betrieben wird, der sein Verhalten zu jedem Zeitpunkt in Hinblick

auf sein Ziel auswählt. Die Agenten melden ihre Werte an das *System*, welches diese in einem *Modell* zusammenführt, was dann weiter zur Entscheidungsfindung genutzt wird. Das System publiziert die Regeln, nach denen es die Messungen verwendet und die Agenten für diese belohnt.

Beginnen wir mit einem ganz einfachem Beispiel: der Agent meldet einen Messwert x, und das System kann den wahren Wert z später feststellen. Das System teilt dem Agenten mit, dass die Messung nach folgender Regel belohnt wird:

$$pay(x,z) = \begin{cases} 1 & \text{falls } x = z \\ 0 & \text{anderweitig} \end{cases}$$

Welches Verhalten wird diese Regel beim Agenten hervorrufen?

Bei korrektem Verhalten kann der Agent eine Belohung von $p(x_1)$ erwarten, da dies die Wahrscheinlichkeit darstellt, dass x_1 gleich dem wahren Wert ist. Nehmen wir an, dass der korrekte Betrieb des Sensors pro Messung Kosten von m verursacht. Bei korrektem Verhalten erwartet der Agent daher einen Gewinn von $p(x_1) - m$. Wenn nun die Posteriori-Wahrscheinlichkeit nur wenig von der Priori-Wahrscheinlichkeit abweicht ($p \simeq q$), dann wird sich der Agent vielleicht eher dafür entscheiden, die Kosten für die Messung einzusparen und einfach den Wert x_0 zu melden. Um dieses zu vermeiden, muss die Belohnung so skaliert werden, dass:

$$m < \underbrace{p(x_1)}_{E_{post}[pay]} - \underbrace{q(x_0)}_{E_{prior}[pay]}$$

Das heisst, anstatt einer Belohnung von 1 sollte eine Belohnung von α versprochen werden, wobei $\alpha \geq \frac{m}{p(x_1)-q(x_0)}$ sein muss. Hiermit würde sich dann ein rationaler Agent für die kooperative Strategie entscheiden.

Es gäbe aber bei einer solchen Regel noch ein anderes Problem. Es erwartet nämlich der Agent auch dann einen Gewinn $\alpha q(x_0)$, wenn er nur nach der Priori-Wahrscheinlichkeit den Wert x_0 meldet. Somit würden wir uns nicht gegen einen Ansturm von Agenten wehren können, die jeder mit geringen Aufwand für fiktive Messungen eine Belohnung kassieren! Wir müssen daher die Regel noch etwas korrigieren, indem wir diesen Erwartungswert abziehen:

$$pay(x,t) = -q(x_0)\alpha + \begin{cases} \alpha & \text{if } x = t \\ 0 & \text{otherwise} \end{cases}$$

Die Schwierigkeit besteht natürlich darin, die richtigen Werte for $q(x_0)$ und α zu bestimmen. Man kann hierfür gewisse Abschätzungen machen: wenn es k mögliche Messwerte gibt, so ist $q(x_0)$ sicher grösser als $1/k$, und α hängt von den Kosten und der Präzision der verfügbaren Sensoren ab. Es wird

aber wohl immmer einiger Experimente bedürfen, um die besten Werte zu ermitteln. Wir werden später einen besseren Mechanismus zeigen, bei dem $\alpha q(x_0)$ nicht benötigt wird.

Wenn die Werte kontinuierliche Zahlen sind, lässt sich das Kriterium der exakten Gleichheit auch auf eine ungefähre Gleichheit erweitern. Zum Beispiel kann man verlangen, dass die Abweichung der Werte unter einem gewissen Grenzwert liegt, oder die Belohnung direkt von der Differenz abhängig machen. Für verifizierbare Messungen ändert dies nichts Wesentliches an den Mechanismen; bei nicht verifizierbaren Messungen bedingt es aber aufgrund des spieltheoretischen Charakters wesentlich tiefere Änderungen, die noch nicht genau erforscht sind.

Oft besteht eine Vorhersage nicht nur aus einem bestimmten Wert, sondern einer Wahrscheinlichkeitsverteilung über alle möglichen Werte. Auch für diese kann man einen Anreizmechanismus entwickeln, indem man die gemeldete Verteilung A in Bezug auf die Grundwahrheit g durch eine *scoring rule* ([13]) evaluiert. Es gibt mehrere solche Scoring Rules, z.B. die Logarithmische:

$$pay(A, g) = C + \ln A(g) \tag{1}$$

wobei C so gewählt werden muss, dass eine positive Belohnung gewährleistet ist.

Nehmen wir zum Beispiel an, dass wir das Wetter für den nächsten Sonntag vorhersagen: wird es regnen, bewölkt oder sonnig sein? Generell mag die Erwartung zum Beispiel so verteilt sein:

$$Pr(P) = \begin{array}{c|c|c} Regen & Wolken & Sonnig \\ \hline 0.2 & 0.3 & 0.5 \end{array}$$

Nach genauer Untersuchung der Wetterlage kommt der Agent nun vielleicht zum Schluss, dass folgende Wahrscheinlichkeitsverteilung vorliegt:

$$Pr(P|o) = \begin{array}{c|c|c} Regen & Wolken & Sonnig \\ \hline 0.8 & 0.15 & 0.05 \end{array}$$

und teilt diese dem System mit. Nehmen wir an, dass es am Sonntag regnet, so belohnt das System den Agenten mit $C + \ln 0.8 = C - 0.22$, wobei dieser allerdings bis zum Sonntag auf die Bezahlung warten muss.

Wieso bietet diese Regel den Anreiz, diese Verteilung richtig zu melden? Der Erwartungswert der Belohnung beträgt bei Verwendung der obengenannten logarithmischen Regel (1):

$$E[pay(A, g)] = \sum_{x \in P} Pr(x|o) pay(A, x) = \sum_{x \in P} Pr(x|o)[C + \ln(A(x)]$$

und somit beträgt der Unterschied zwischen korrektem und unkorrektem Bericht:

$$\begin{aligned} E[(pay(Pr,g)] & -E[pay(A,g)] \\ = & \sum_{x\in P} Pr(x|o)[(C+\ln Pr(x)) - (C+\ln A(x))] \\ = & \sum_{x\in P} Pr(x|o)\ln\frac{Pr(x)}{A(x)} \\ = & D_{KL}(Pr||A) \end{aligned}$$

Da die Gibbsche Ungleichung besagt, dass die Kullback-Leibler Divergenz $D_{KL}(Pr||A) \geq 0$ sein muss, ergibt die Strategie, korrekt $A = Pr$ zu melden, die grösste Belohnung. Der Trick liegt hier also darin, dass die wirkliche Meinung des Agentent in die Berechnung des Erwartungswertes der Belohnung einfliesst und man daher diesen Erwartungswert genau für die aufrichtige Meldung maximiert.

3.2 Reputation

In gewissen Fällen haben wir es mit Agenten zu tun, die auf Anreize nicht rational reagieren. Dieses passiert im Wesentlichen in folgenden Fällen:

- defekte oder falsch kalibrierte Sensoren, oder
- mutwillige Manipulation zum Vertuschen der wirklichen Daten (z.B. durch den Verschmutzer selbst).

In diesem Fall sind Anreize nutzlos. Wir können aber auf den Leumund (Reputation) von Agenten zurückgreifen, um trotzdem schlechte Daten zu eliminieren. Diese Idee ist nicht neu und es gibt verschiedene Mechanismen, die für Sensoren implementiert wurden. Das bekannteste ist wohl das β-Reputationssystem [14]. Hier wird jedem Agenten eine Zahl zugeordnet, die den Leumund darstellt und sich wie folgt ergibt:

$$Rep = \frac{|\{korrekteMessungen\}|}{|\{alleMessungen\}|}$$

wobei man jeweils durch Vergleich mit dem wirklichen Wert feststellt, ob eine Messung korrekt war oder nicht.

Man benützt dann diese Zahl, um die zu erwartende Qualität eines Messwertes zu begutachten. Normalerweise zieht man für das Modell nur Messwerte von Agenten in Betracht, deren Leumund einen gewissen Mindestwert

überschreitet. Hierdurch werden mit einer gewissen Verzögerung diejenigen Agenten herausgefiltert, die schlechte Messwerte liefern, und man hält so sowohl böswillige Agenten als auch fehlerhafte Sensoren vom System fern.

Es stellt sich allerdings heraus, dass solch einfache Methoden in Wirklichkeit keinen Schutz gegen bösartige Agenten bilden. Es gibt nämlich einfache Strategien, mit denen diese sie umgehen können. Zum Beispiel kann ein Agent zunächst durch Lieferungen von korrekten Daten, die mit anderen Sensoren redundant sind, einen guten Leumund erzielen. Dann kann dieser Leumund später dazu benutzt werden, mit falschen Daten die Entscheidungen zu beeinflussen. Da sich immer wieder herausstellte, dass sie gegen derartige Strategien nicht resistent sind, haben sich Reputationssysteme bis heute nicht durchgesetzt.

Wir haben aber kürzlich eine Lösung zu diesem Problem gefunden, die es erlaubt, den Einfluss von bösartigen Sensoren auf die Entscheidungen beliebig klein zu halten und das Einhalten dieser Grenze auch unabhängig von der Strategie der Agenten zu garantieren. Die wesentliche Erkenntnis, die das erlaubt, ist dass man nicht die gemeldeten Werte selber, sondern deren Einfluss auf das daraus konstruierte Modell betrachtet. Hiermit kann man nämlich einflussreiche von einflusslosen Meldungen unterscheiden, und somit obengenannte Manipulationen verunmöglichen. Wir nennen diese Technik den Einflussbegrenzer (Influence Limiter) ([15]).

Sie funktioniert wie folgt. Wir nehmen allgemein an, dass das System alle gemeldeten Messwerte in ein Modell integriert, welches dann weiter zur Entscheidungsfindung genutzt wird. Das Modell erlaubt es insbesondere, für beliebige Punkte den Messwert als Wahrscheinlichkeitsverteilung vorherzusagen. Für jeden gemeldeten Messwert wird der Einfluss auf das Modell evaluiert. Diese Evaluation bezieht sich auf die Vorhersage eines Referenzpunktes x, für den eine präzise Messung z vorliegt.

Für einen gemeldeten Wert a vergleichen wir also die Vorhersagen $Pr_{-a}(x)$ ohne Berücksichtigung von a und $Pr(x)$ mit Berücksichtigung desselben. Für die Beurteilung der Qualität dieser Verteilungen verwenden wir Scoring Rules, wie wir sie schon oben für die Eruierung von Wahrscheinlichkeitsverteilungen vorgeschlagen haben. Wir berechnen also für die Meldung x mithilfe einer Scoring Rule S eine Bewertung

$$s(x) = S(Pr, z) - S(Pr_{-x}, z)$$

Diese Bewertung kann sowohl positiv als auch negativ ausfallen, und gibt den Einfluss der gemeldeten Messung auf das Modell wieder.

Das Ziel des Einflussbegrenzers ist es, die Summe der Einflüsse aller

von einem Agenten eingereichten Messungen von unten zu begrenzen, d.h. negative Einflüsse nur bis zu einem gewissen Grenzwert zuzulassen.

Wir beschreiben diese Einflüsse durch den Leumund (reputation) ρ. Dieser ist am Anfang ρ_0 und wird nach jeder Meldung in folgender Art angepasst:

$$\rho_{posterior} = \rho_{prior}(1 + \frac{1}{2}s(x))$$

Wir zeigen in [15], dass hierdurch die Summe der Einflüsse von unten durch $-2\rho_0$ begrenzt wird, d.h. dass sich die schädlichen Einflüsse eines einzelnen Sensors nicht auf mehr als $2\rho_0$ summieren können. Hiermit haben wir zum ersten Mal eine Garantie, dass unabhängig von der Strategie eines böswilligen Agenten der Schaden in Grenzen gehalten werden kann. Hiermit kann also das Problem erstmals wirklich gelöst werden.

Neben der Begrenzung von negativen Einflüssen ist es aber auch wichtig, die positiven Einflüsse nicht zu unterdrücken. Man könnte ja einfach alle Meldungen ignorieren, und damit hätte man sicher auch alle falschen Werte eliminiert! Das zweite wichtige Resultat der Arbeit in [15] ist daher, dass wir genügend Information behalten, damit das Gesamtsystem das no-regret Kriterium erfüllt: die Summe der Fehler des Modells ist beschränkt, und der durchschnittliche Fehler strebt somit gegen 0.

Wir zeigen später in diesem Beitrag, dass wir in der Praxis wesentlich besseres Verhalten erwarten können, als es durch diese Grenzwerte gegeben ist.

4 Unverfizierbare Messungen

In vielen Fällen ist es allerdings auch im Nachhinein nicht möglich, die Messwerte zu verifizieren. In diesem Fall könnten sich allerdings die verschiedenen Messungen gegenseitig verifizieren. Der Mechanismus wird hiermit zu einem *Spiel*, in dem die Kombination der Strategien der Agenten die Belohnung bestimmt. Die einfache Auswahl einer Strategie durch die Optimierung der erwarteten Belohnung wird hier ersetzt durch die Wahl eines spieltheoretischen Gleichgewichts, in der die Strategie jedes Agenten die bestmögliche Antwort auf die Strategien der anderen Agenten ist.

Für die gegenseitige Verifizierung ist es nötig, für jeden Bericht einen *Peer*-Agenten zu bestimmen, der dasselbe messen sollte. Am besten findet man natürlich einen Sensor am gleichen Ort. Da aber ein solcher normalerweise nicht existiert, behilft man sich meistens damit, dass man durch Interpolation von benachbarten Messwerten einen künstlichen Peer erzeugt und

diesen dann als Grundwahrheit einsetzt.

Als Beispiel nehmen wir einmal an, die Messung ergibt einen von 3 möglichen Werten: a, b oder c. Wenn man die Belohung durch die einfache Regel 3.1 berechnet, erhält man damit damit ein einfaches Spiel, was auch *output agreement* genannt wird:

		peer		
		a	b	c
agent	a	(1,1)	(0,0)	(0,0)
	b	(0,0)	(1,1)	(0,0)
	c	(0,0)	(0,0)	(1,1)

wobei die Einträge der Matrix die Belohnung für (Agenten,Peer) darstellen.

Solange der Agent glaubt, dass der Peer eine korrekte Messung meldet, stellt die korrekte Messung ein Gleichgewicht dar: der Agent erwartet, dass der Peer eine kooperative Strategie verfolgt und den echten Wert meldet, und damit muss er auch denselben Wert finden.

Allerdings hat dieses Spiel auch andere Gleichgewichte. Zum Beispiel könnten alle Agenten immer den Wert a melden - dann wäre ebenfalls die maximal mögliche Belohnung gesichert. Daher lässt sich diese einfache Regel nicht so leicht anwenden.

Ein weiteres Problem ergibt sich, wenn die Agenten nicht sicher sind, dass ihr Peer genau denselben Wert misst. Was wäre, wenn mein Sensor aufgrund eines Feuers sehr hohe Luftverschmutzung misst? Die Interpolation der benachbarten Messungen zeigt diese vielleicht nicht so stark, so dass ich vielleicht eher einen abgeschwächten Wert melden sollte. Hier muss das erhöhte Risiko einer aufrichtigen Meldung durch eine höhere Belohnung kompensiert werden.

Ein besseres Prinzip ist es, die Anreize aufgrund der *Veränderungen* des Glaubens anstatt des Glaubens selber zu berechnen. So werde ich beim Beispiel eines Feuers zwar vielleicht nicht glauben, dass mein Peer einen ähnlich hohen Messwert meldet, aber ich werde dem doch eine erhöhte Wahrscheinlichkeit zuschreiben. Im Allgemeinen kann man annehmen, dass sich die Wahrscheinlichkeit, die ich meiner wirklichen Messung zuschreibe, am meisten von allen Werten erhöht.

Auf dieser Grundlage können wir nun eine bessere Regel definieren, die wir *Peer Truth Serum* (Peer Wahrheitsserum) nennen. Sie nimmt an, dass für einen bestimmten Messpunkt X eine Priori-Wahrscheinlichkeitsverteilung $R(X)$ vorliegt, die sowohl dem System wie auch den Agenten bekannt ist. $R(X)$ könnte z.B. durch die bereits vorliegenden Messungen geschätzt sein.

Wir benutzen nun folgende Regel, um die Belohnung für die Meldung x_i zu berechnen, wenn der Peer den Wert x_j gemeldet hat:

$$Pay(x_i, x_p) = \begin{cases} 1/R(x_i) & \text{if } x_i = x_p \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Wir wollen nun herleiten, unter welchen Bedingungen diese Regel einen Agenten dazu bewegt, eine genaue Messung zu melden. Wenn wir die Priori-Wahrscheinlichkeitsverteilung des Agenten mit Pr bezeichnen, und mit Pr_i die Posteriori-Verteilung nach Messung des Wertes x_i, so erhalten wir für die Bedingung:

$$\begin{aligned} E_{Pr_i(r)}[Pay(x_i, r)] &> E_{Pr_i(r)}[Pay(x_j, r)] \\ Pr_i(x_i)Pay(x_i, x_i) &> Pr_i(x_j)Pay(x_j, x_j) \\ Pr_i(x_i)/R(x_i) &> Pr_(x_j)/R(x_j) \end{aligned}$$

Wenn wir weiter annehmen, das die Verteilungen R und Pr gleich sind, erhalten wir die folgende Bedingung, die wir *self-predicting* nennen:

$$\frac{P_i(x_i)}{P(x_i)} > \frac{P_i(x_j)}{P(x_j)} \text{ sofern } i \neq j \tag{3}$$

was gleichbedeutend damit ist, dass die Posteriori-Verteilung Pr_i so gewählt ist, dass der Messwert x_i die Maximum-Likelihood Schätzung ist. Da diese Bedingung der gängigen Praxis der Signalverarbeitungentspricht, kann man sie weitgehend als gegeben annehmen.

Somit schafft das Peer Truth Serum also sehr weitgehend den Anreiz zur wahrheitsgetreuen Messung, sofern die Verteilung R richtig gesetzt ist.

Um die richtige Verteilung R zu finden, kann man sich dadurch behelfen, dass man R direkt aus den gemeldeten Messwerten herleitet. Wie in der Abbildung 2 gezeigt, leitet man diese aus der Gesamtheit der Messungen her. Wenn der Agent nichts genaues über die Messung weiss, ist die Annahme gerechtfertigt, dass diese wie R verteilt ist. Im Vergleich damit sollten aber die Peers, die ja in der Nachbarschaft des Messpunktes angesiedelt sind, einen wesentlich ähnlicheren Messwert aufweisen.

Am einfachsten kann man nun den Peer zufällig aus der Nachbarschaft auswählen und dann die Regel 2 mit dem von diesem erhaltenen Messwert anwenden. Durch die zufällige Auswahl hat jedoch die Belohnung eine hohe Varianz. Diese kann man reduzieren, indem man den Durchschnitt über alle in Frage kommenden Peers bildet. Damit erhält man folgende, recht einfache Formel:

$$Pay(x_i, P) = \frac{P_{local}(x_i)}{P_{global}(x_i)}$$

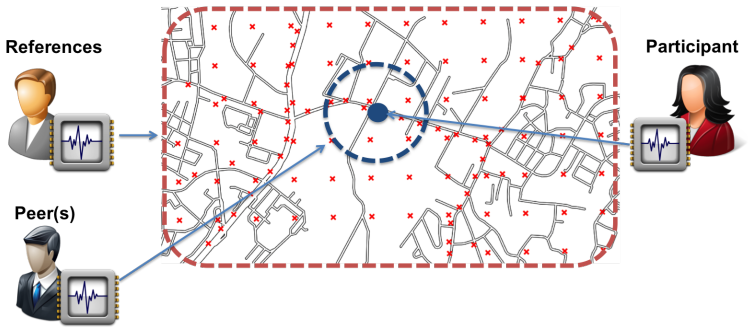


Figure 2: Szenario für Crowdsensing von Luftverschmutzung. Die Priori-Verteilung R wird aus der Gesamtheit der Messwerte bestimmt. Peers werden aus der Nachbarschaft ausgewählt.

wobei:

- $P_{local}(x_i)$ = Anteil der peers, die x_i melden, und
- $P_{global}(x_i)$ = Anteil der x_i im gesamten Messbereich.

Wie bei unserem einfachsten Mechanismus 3.1 müssen wir noch eine Konstante abziehen, so dass zufällige Meldungen keine Belohnung erhalten, sowie die Bezahlung skalieren, so dass sie die Kosten für die Messung abdeckt.

Wir erhalten damit die Funktion des *Peer Truth Serum for Crowdsourcing* (PTSC):

$$Pay(x_i, P) = \alpha \left(\frac{P_{local}(x_i)}{P_{global}(x_i)} - 1 \right) \tag{4}$$

Hier muss jetzt nur noch die Konstante α hinreichend gross gewählt werden, um die Kosten der Messung zu decken.

Das PTSC hat mehrere wünschenswerten Eigenschaften. Die Agenten müssen hier überhaupt keine Priori-Wahrscheinlichkeitsverteilung über die Messwerte haben. Weiterhin hat das Spiel zwischen den Agenten ein Gleichgewicht, in dem alle Agenten genaue Messungen abliefern, sofern diese die Bedingung 3 erfüllen. Andere Gleichtgewichts-Strategien ergeben eine geringere Belohnung, sofern sie nicht einfach eine Permutation der exakten Strategien sind (alle Agenten berichten vertauschte Werte). Ausserdem ist die erwartete Belohnung für zufällige Meldungen nach der Verteilung P_{global} gleich null.

Es lässt sich somit sehr gut in der Praxis anwenden, was auch in mehreren Studien [16] gezeigt wurde.

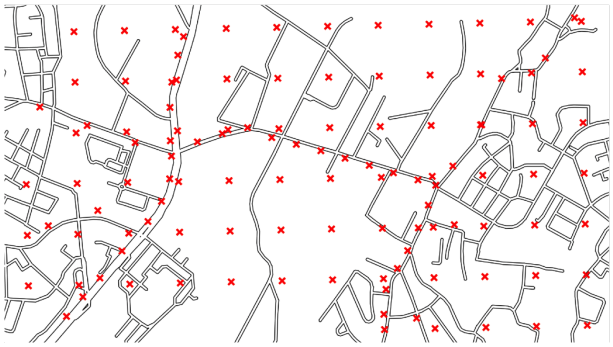


Figure 3: Der Bezirk von Strassbourg, der für die Simulation herangezogen wurde. Die Messpunkte sind in Rot eingezeichnet.

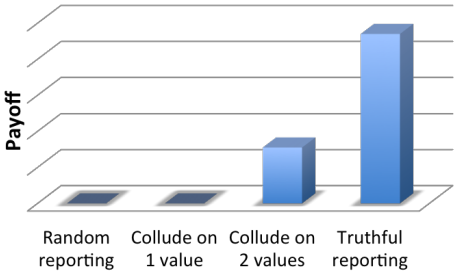


Figure 4: Durchschnittliche Belohnung für verschiedene Strategien.

5 Praktisches Beispiel

Wir zeigen die Anwendung der hier gezeigten Mechanismen an Hand einer Simulation des ADMS Modells [17], die vor einigen Jahren von Umweltforschern für einen Teil der Stadt Strassbourg mithilfe von Sensordaten erstellt wurde und diese recht genau abbildet. In diesem Bereich haben wir für eine Mischung aus wirklichen Sensoren und durch das Modell simulierte Sensoren kontinuierlich über 3 Wochen die Daten der NO_2 Konzentrationen aufgenommen und die Messwerte auf 4 Wertebereiche diskretisiert. Die Abbildung 3 zeigt diesen Bereich.

5.1 Peer Truth Serum

Wir zeigen zunächst Simulationen des Peer Truth Serum for Crowdsourcing (PTSC, 4). Abbildung 4 zeigt, wie sehr die durchschnittliche Belohnung

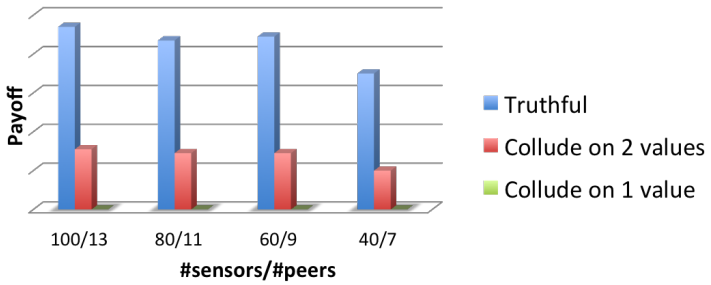


Figure 5: Durchschnittliche Belohnung für verschiedene Anzahlen von Sensoren und Peers.

von verschiedenen Strategien abhängt, und wie die Funktion die Lieferung korrekter Daten belohnt. Es werden die folgenden Strategien verglichen:

1. Random reporting: Meldung zufälliger Werte gemäss der Priori-Verteilung.
2. Collude on 1 value: alle Agenten melden immer denselben Wert.
3. Collude on 2 values: alle Agenten, die einen der 2 niedrigeren Werte messen, berichten den niedrigsten Wert; alle, die einen der 2 höheren Werte messen, den höchsten Wert (somit eine Reduktion der Auflösung).
4. Truthful reporting: alle Agenten verfolgen die kooperative Strategie.

Man sieht deutlich, das die kooperative Strategie deutlich besser als die anderen Strategien abschneiden; wir bemerken, dass die einzige Strategie mit nennenswerter Belohnung ebenfalls eine Messung (wenn auch ungenauer) vorraussetzt.

Abbildung 5 zeigt ferner, dass diese Methode trotz der stärker ins Gewicht fallenden Approximationsfehler auch für wenige Sensoren noch gilt.

Es ist auch interessant zu sehen, wie sehr die Belohnung einzelner Sensoren variiert. Wie in Abbildung 6 zu sehen ist, gibt es für die Sensoren recht unterschiedliche Möglichkeiten. Ein Sensor in einem Gebiet, in dem starke Schwankungen auftreten, kann nämlich mit sehr viel mehr Belohnung rechnen als ein Sensor, der nur wenige Schwankungen zu melden hat. Das heisst auch, dass es attraktiv ist, in Gebieten mit starken Schwankungen einen Sensor zu betreiben, so dass sich dort auch mehr Messwerte finden werden.

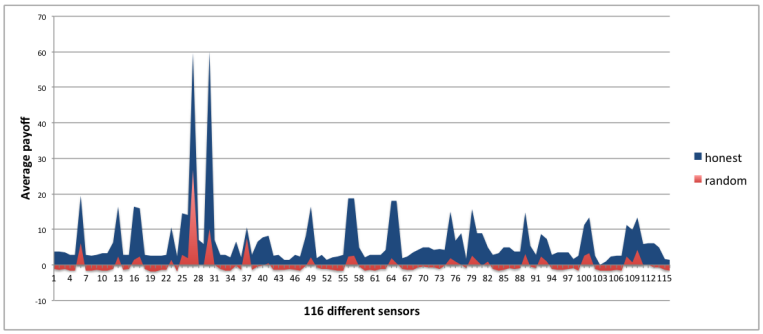


Figure 6: Kumulative Belohnung der einzelnen Sensoren während der gesamten Simulation.

Man kann in Abbildung 6 auch gut sehen, dass die korrekten Strategien (in Blau) wesentlich höhere Belohnungen also zufällige Meldungen (in Rot) versprechen, und zwar ausnahmslos für alle Sensoren in der Simulation.

5.2 Einflussbegrenzer

Wir benutzen dasselbe Simulationsmodell, um die Funktion des Einflussbegrenzers zu illustrieren. Wir simulieren hierbei vier verschiedene bösartige Strategien der Agenten:

1. vary: zuerst wird mit 1000 korrekten Messwerten die Reputation aufgebaut, und sodann fortwährend mit sehr niedrigen Werten versucht, das Modell zu verfälschen.

2. deceive: die oben schon angesprochene Strategie, mit korrekten Werten die Reputation aufzubauen, solange diese unter einem Grenzwert liegt, und immer dann falsche sehr niedrige Werte einzuschleusen, wenn die Reputation darüber liegt.

3. vary and deceive: zuerst vary und dann deceive.

4. cover: ähnlich wie deceive, aber berichtet falsche Werte nur dann, wenn die korrekten Werte zu hoch sind.

Abbildung 7 zeigt den Modellfehler in der Simulation mit 15 ehrlichen und 30 bösartigen Sensoren, die eine der 4 obengenannten Strategien verfolgen. Untersucht wird der Einflussbrenzer (CSIL) gegen das Beta-System,

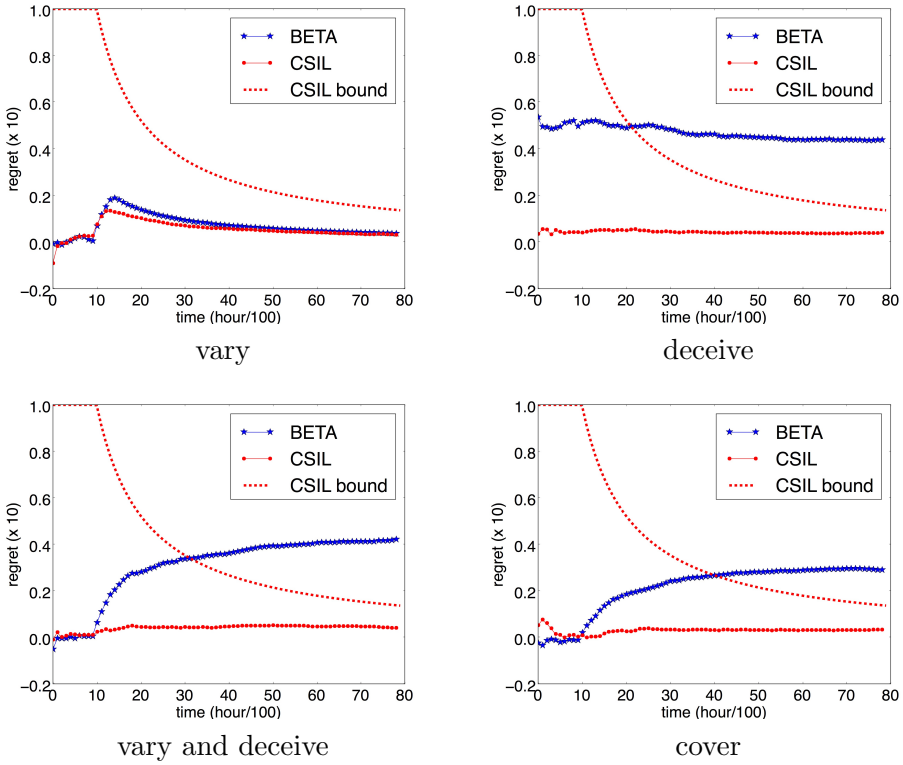


Figure 7: Einflussbegrenzer: Bedauern für verschiedene Absprachestrategien.

was oben erwähnt wurde. Wir zeigen auch den Verlauf des für den Einflussbegrenzer gültigen Grenzwertes (CSIL bound). Man sieht deutlich, dass sowohl Beta als auch CSIL die vary-Strategie gut abwehren können. Anders sieht es allerdings bei den 3 anderen Strategien aus, gegen die BETA nicht viel ausrichten kann, während der Einflussbegrenzer fast überhaupt keine Einwirkung zulässt. Man sieht auch, dass das wirkliche Verhalten des Einflussbegrenzers noch um einiges besser ist, als es die theoretische Grenze erwarten liesse.

6 Diskussion

Um die Vision von Smart Cities Wirklichkeit werden zu lassen, ist eine kontinuierliche Sammlung von Daten unerlässlich. Wir haben gezeigt, wie es möglich ist, diese Daten durch andere Agenten erheben zu lassen, solange die Agenten ein rationales Verhalten an den Tag legen.

Wir haben zwei neuartige Mechanismen vorgestellt. Der Einflussbegrenzer (Influence Limiter) erlaubt es uns, den Einfluss von bösartigen oder fehlerhaften Messwerten auf das Gesamtmodell zu begrenzen. Es handelt sich hier um ein neuartiges Reputationssystem, was die Manipulierbarkeit früherer Systeme beweisbar vermeidet.

Das Peer Truth Serum erlaubt es uns, für Daten zu bezahlen, ohne dass damit auch inkorrekte Daten belohnt werden. Es hat auch den weiteren Vorzug, dass Sensoren in Gebieten mit stark schwankenden Werten wesentlich höhere Belohnungen erzielen können. Hiermit wird sich die selfselection dahingehend auswirken, dass solche Bereiche auch durch eine grössere Anzahl von Sensoren abgedeckt werden. Wir verweisen auch auf frühere Arbeiten mit ähnlichen Zielen, so etwa *Peer Prediction* ([18]) und das *Bayesian Truth Serum* [19] mit den Verfeinerungen in [20] und [21].

Die Forschung auf dem Gebiet von Crowdsourcing-Mechanismen steckt heute noch in den Kinderschuhen. Die Mechanismen sind noch nicht für kontinuierliche Werte ausgelegt, obwohl dazu gewisse Ansätze bestehen ([22]). Es werden auch nur einfache Daten erhoben, und keine komplexen Zusammenhänge. Wir benötigen auch eine gewisse Homogenität der Sensoren, damit für alle Sensoren die gleichen Anreize gelten können. Bis ist heute keiner der Mechanismen praktisch umgesetzt worden, da sie ja erst in den letzten Jahren entdeckt wurden. Wir erhoffen uns der Praxis weitere Erkenntnisse zur Forschung.

References

[1] M. Eeftens et al., Spatial variation of PM2.5, PM10, PM2.5 absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO2 Results of the ESCAPE project, Atmospheric Environment **62**, pp. 303-317, 2012.

[2] J. Cyrys et al., Variation of NO2 and NOx concentrations between and within 36 European study areas: Results from the ESCAPE study, Atmospheric Environment **62**, pp. 374-390, 2012.

[3] `http://www.waze.com/` Retrieved 2016-05-05

[4] C. Ravetti, Y. Jin, M. Quan, Z. Shiqiu, T. Swanson. A Dragon eating its own tail: public information about pollution in China, Paper 27, Graduate Institue (IHEID), Geneva, September 2015.

[5] M. Müller, Pekings Propagandisten sehen rot, Neue Zürcher Zeitung, 14. Dezember 2015.

[6] M.D. Müller, D. Hasenfratz, O. Saukh, M. Fierz, C. Hueglin. Statistical modelling of particle number concentration in Zürich at high spatiotemporal resolution utilizing data from a mobile sensor network. Atmospheric Environment, 126, pp. 171-181, 2016.

[7] A. Jutzeler, J.J. Li and B. Faltings. A Region-Based Model for Estimating Urban Air Pollution. Proceedings of the 28th conference of the AAAI, 424-430, 2014.

[8] Kickstarter: "#AirQualityEgg by #SenseMakers". kickstarter.com. Retrieved 2014-08-23

[9] A. Vaughan: "Pigeon patrol takes flight to tackle London's air pollution crisis," The Guardian, March 14th, 2016.

[10] Origins Technology: Laser Egg. Retrieved 2016-05-05.

[11] A. Papakonstantinou, A. Rogers, E. H. Gerding, N. R. Jennings, Mechanism design for the truthful elicitation of costly probabilistic estimates in distributed information systems, *Artificial Intelligence* **175**(2), pp. 648–672, 2011.

[12] B. Faltings, J.J. Li, R. Jurca, Incentive Mechanisms for Community Sensing. *IEEE Transaction on Computers* **63**(1), pp. 115–128, 2014.

[13] L.J. Savage. Elicitation of Personal Probabilities and Expectations. *Journal of the American Statistical Association*, **66**(336),pp. 783–801, 1971

[14] A. Jøsang , R. Ismail: "The beta reputation system," Proceedings of the 15th Bled Conference on Electronic Commerce, 2002.

[15] G. Radanovic and B. Faltings. Limiting the Influence of Low Quality Information in Community Sensing. Proceedings of the 15th international conference on autonomous agents and multiagent systems (AAMAS'16), 2016.

[16] G. Radanovic, B. Faltings and R. Jurca. Incentives for Effort in Crowdsourcing Using the Peer Truth Serum. ACM Transactions on Intelligent Systems and Technology (TIST), **7**(48), 2016.

[17] R. N. Colvile, N. K. Woodfield, D. J. Carruthers, B. E. A. Fisher, A. Rickard, S. Neville, and A. Hughes. Uncertainty in dispersion modelling and urban air quality mapping. Environmental Science and Policy, **5**(3), pp. 207–220, 2002.

[18] N. Miller, P. Resnick, R. Zeckhauser, Eliciting Informative Feedback: The Peer-Prediction Method. *Management Science* **51**, pp. 1359–1373, 2005.

[19] D. Prelec, A Bayesian Truth Serum for Subjective Data. *Science* **306**(5695), pp. 462–466, 2004.

[20] J. Witkowski, D.C. Parkes, A robust bayesian truth serum for small populations. *Proceedings of the 26th AAAI Conference on Artificial Intelligence*, pp. 1492–1498, 2012.

[21] G. Radanovic, B. Faltings. A Robust Bayesian Truth Serum for Non-binary Signals. *Proceedings of the 27th AAAI Conference on Artificial Intelligence*, pp. 833–839, 2013.

[22] G. Radanovic and B. Faltings. Incentives for Truthful Information Elicitation of Continuous Signals. In Proceedings of the 28th AAAI Conference on Artificial Intelligence, pp. 770-776, 2014.