

Datasets

Para a aplicação e teste da abordagem proposta é necessário ter um conjunto de dados com atividades realizadas por usuários em ambientes IoT. Uma busca na literatura foi realizada com intuito de encontrar base de dados abertas que possuísem as informações necessárias para a criação do perfil do usuário. Como resultado dessa busca foi encontrado o Projeto CASAS <<http://casas.wsu.edu/>>, que contém diversos *datasets* com informações sobre atividades realizadas por moradores em suas casas.

Os *datasets* estudados foram divididos em quatro categorias: i) 3 *Datasets* na cidade de Kyoto com informações de dois moradores de um apartamento; ii) 2 *Datasets* na cidade de Tulum com informações de um casal; iii) *Dataset* na cidade de Aruba contendo as informações de uma moradora em seu apartamento; iv) *Dataset* contendo as informações de uma pessoa em seu apartamento.

1. WSU Kyoto Smart Apartment - Two Resident Testbed

Descrição: Três *datasets* com atividades diárias de dois moradores na mesma residência em várias datas diferentes. A disposição dos sensores usados para a formação dos *datasets* é dividida em duas etapas. A primeira, sem anotação, já a segunda utilizando de anotação dos tipos de sensores. Para este trabalho, foi-se observado o esquema dos sensores com anotação.

Datasets:

- WSU Smart Apartment 2010 Two Resident Testbed (Dataset 11)
- WSU Smart Apartment Summer 2009 Two Resident Testbed (Dataset 8)
- WSU Smart Apartment Spring 2009 Two Resident Testbed (Dataset 7)

Categorização dos sensores: Os autores categorizam os sensores em oito tipos diferentes.

- **M:** Sensores de movimento;
- **I:** Itens selecionados na cozinha;
- **D:** Sensores de porta;
- **AD1-A:** Sensores de queimadura;
- **AD1-B:** Sensores de água quente;
- **AD1-C:** Sensores de água gelada;
- **T:** Sensores de temperatura;
- **P001:** Sensores de uso de eletricidade.

Formato da casa: A casa utilizada no experimento é dividida em dois andares.

- Primeiro andar: sala, cozinha, lavanderia e escada.
- Segundo andar: três quartos, um corredor e um banheiro.

Disposição dos sensores no ambiente:

Sala (L008):

- **Porta de Entrada:** Sensor de Porta D002;
- **Da porta de Entrada a Porta direita da cozinha:** Sensores de movimentação M012 - M015;
- **Espalhados pelo meio da sala:** sensores de movimentação M002 - M011;
- **No meio da sala:** Entre os sensores de movimentação M009 e M010 tem dois sensores de itens selecionados na cozinha I008 e I009;
- **No meio da sala:** Junto ao sensor M006 existe um sensor de temperatura (T001);
- **Canto inferior esquerdo:** Temos dois sensores de item na cozinha I011 e I012 (de acordo com o primeiro modelo, seriam itens de uma TV), um sensor de porta D013, além disso, temos o sensor do usuário 2 (R002) ao lado do sensor M002;
- **Porta esquerda de entrada para a cozinha:** Sensor de movimento M001;
- **Obs.:** Existe um sensor E001, porém os autores não deram um *label* para esse sensor, portanto não foi possível identificar o tipo do sensor.

Cozinha (L009 - L011): A cozinha é dividida em três etapas: i) da porta de acesso a sala pela esquerda até a porta de entrada no canto inferior esquerdo da cozinha; ii) uma espécie de lavanderia; iii) corredor direito que dá acesso a sala pela porta direita. Além disso, temos uma pia e um freezer.

- **(L009) Entrada a esquerda para a sala até o canto inferior esquerdo da cozinha:** Sensores de movimentação M023, M022, M021 e M024. Sensor de porta D001 ligado a porta que dá acesso a cozinha pelo canto inferior esquerdo. Sensor de movimentação M025;
- **(L010) Entrada a direita para a sala até o canto inferior direito da cozinha:** Sensores de movimentação M016 - M018, colado ao sensor M017 temos um sensor de temperatura T002;
- **(L010) Canto inferior direito (Espécie de dispensa):** Sensor de movimento M051 e um sensor de porta D011;
- **(L010) Pia:** No canto superior tem-se um sensor de porta D016, ao meio temos um sensor de porta D007 e um item de cozinha I007 (de acordo com o primeiro modelo, esses sensores monitoram o *Burner*). No canto inferior tem-se dois sensores de porta D014 e D015 (torneiras), logo acima das torneiras existem três sensores A001 (Stove), A002 (água quente) e A003 (água gelada);
- **(L010) Prateleira:** Por cima da pia existe uma prateleira com itens de cozinha I001 - I006 e I010.
- **(L011) Freezer (Meio da cozinha) :** Sensor de porta D012 para abrir o freezer, sensores de porta D008, D009 e D010 (No primeiro modelo, relacionados à geladeira e microondas);
- **(L011) Lavanderia:** Sensores de Movimentação M019 e M020;
- **Escada:** Sensor de movimentação M026.

Corredor (L003):

- **Da escada até a porta do banheiro:** Sensores de movimentação M027 - M029.

Quarto 1 (L001):

- **Canto inferior esquerdo até o canto superior esquerdo:** Sensores de movimentação M045 - M047;
- **Canto inferior direito até o canto superior direito:** Sensores de movimentação M050 - M048. Ao lado do sensor M049 existe um sensor de temperatura T004;
- **Porta:** Dois sensores de movimento M043 e M044, além disso, um sensor de porta D004.

Quarto 2 (L002):

- **Entrada do quarto:** Sensor de temperatura T005 e um sensor de movimentação M042.
- **Obs.:** Existe um sensor E002, porém os autores não deram um *label* para esse sensor, portanto não foi possível identificar o tipo do sensor.

Quarto 3 (L004):

- **Canto inferior esquerdo até o canto superior esquerdo:** Sensores de movimentação M033 - M031. Junto ao sensor M031 existe um sensor de temperatura T003;
- **Canto inferior direito até o canto superior direito:** Sensores de movimentação M034 - M036;
- **Porta:** Um sensor de movimento M030, além disso, um sensor de porta D003.

Banheiro (L005 - L007):

- **Porta:** Sensor de movimentação M037 e sensor de porta D005;
- **(L005) Pia:** Sem sensores;
- **(L006) Entrada do banheiro até a porta do sanitário:** Sensores de movimentação M038 - M040;
- **(L007) Área do sanitário:** Sensor de movimentação M041 e sensor de porta D006.
- **Obs.:** O sanitário e a banheira foram marcados com os *labels* F002 e F001. Acredita-se que é para informar que os mesmos não foram utilizados durante o experimento.

Atividades dos Usuários:

Para o dataset mais recente (WSU Smart Apartment 2010 Two Resident Testbed) os autores conseguiram identificar 13 atividades diferentes através dos sensores utilizados:

- Banho;
- Transição da cama para o banheiro;
- Comendo;

- Entrando na casa;
- Limpando a casa;
- Saindo de casa;
- Preparando refeição;
- Higiene pessoal;
- Dormindo;
- Dormindo fora da cama;
- Caminhando no quarto;
- Assistindo tv;
- Trabalho.

Já nos outros *datasets* foram mapeadas apenas 4 atividades por usuário:

- Transição da cama para o banheiro;
- Higiene pessoal;
- Dormindo;
- Trabalho.

Sendo que essas quatro atividades também aparecem no *dataset* mais recente. Além disso, outras atividades foram mapeadas, entretanto, os autores não vinculam a nenhum dos dois usuários, ou seja, não dá para saber quem foi que realizou aquela atividade. A seguir as atividades sem usuários:

- Limpando;
- Preparando comida;
- Estudando;
- Limpando a banheira;
- Assistindo tv.

Estrutura do Arquivo de Dados: O arquivo é dividido em 5 colunas:

- **Data de sensoriamento:** Informa o período em que os dados foram coletados de cada sensor;
- **Tempo:** Em que horário aquele sensor foi utilizado para captar os dados do ambiente;
- **Sensor:** Identifica qual sensor da casa foi utilizado para capturar aquela informação (ex., M046);
- **Informação do sensor:** Informa se o sensor está ligado ou não, além de algumas outras informações que variam dependendo do tipo do sensor;
- **Tags:** Denotando em que momento começa determinada atividade do usuário 1 ou 2.

2. WSU Tulum Smart Apartment - Two Resident Testbed

Descrição: Dois *datasets* com atividades diárias de um casal em sua residência em datas diferentes. Nos dois *datasets* as atividades são anotadas, com marcações mostrando o início e o fim da atividade.

Datasets:

- WSU Tulum Smart Apartment 2009-2010 Two Resident Testbed (Dataset 10)
- WSU Tulum Smart Apartment 2009 Two Resident Testbed (Dataset 9)

Categorização dos sensores: Os autores categorizam os sensores em dois tipos diferentes.

- **M:** Sensores de movimento;
- **T:** Sensores de temperatura.

Existe uma variação no número de sensores de um *dataset* para o outro. No *dataset* mais recente (WSU Tulum Smart Apartment 2009-2010 Two Resident Testbed) existem 31 sensores de movimentação e 5 sensores de temperatura, já no *dataset* mais antigo (WSU Tulum Smart Apartment 2009 Two Resident Testbed) existem 18 sensores de movimentação e 2 de temperatura.

Formato da casa: Os autores não disponibilizam o mapa dos sensores do *dataset* mais antigo, somente do *dataset* mais atual. A casa utilizada no experimento é dividida em dois andares.

- Primeiro andar: terraço, sala, cozinha, dois closets, corredor de entrada e escada.
- Segundo andar: dois quartos, um corredor, um banheiro e dois closets.

Disposição dos sensores no ambiente:

Terraço:

- Nenhum sensor foi instalado no terraço da residência.

Sala:

- **Lado esquerdo:** Dois sensores de movimentação M011 e M012;
- **Meio:** Dois sensores de movimentação M009 e M010;
- **Lado direito:** Dois sensores de movimentação M007 e M008, além disso, um sensor de temperatura T002 ao lado do sensor M008.

Cozinha:

- **Pia:** Sensor de movimentação M016;
- **Entrada da Cozinha até a pia:** Sensores de movimentação M013 - M015, sensor de movimentação M003 próximo a pia. Além disso, sensor de temperatura T001 ao lado do sensor M015.

Closets: Sem sensores instalados.

Corredor de Entrada:

- **Da porta de entrada até a sala:** Três sensores de movimentação M001, M005 e M006;
- **Canto inferior direito:** Sensor de movimentação M004.

Escada: Sensor de movimentação M023.

Corredor:

- **Próximo ao quarto superior:** Sensores de movimentação M024.

Quarto Superior:

- **Canto esquerdo:** Sensores de movimentação M028 e M030;
- **Canto direito:** Sensor de movimentação M029;
- **Porta:** Dois sensores de movimento M017 e M031, além disso, um sensor de temperatura T004.

Quarto Inferior:

- **Canto esquerdo:** Sensores de movimentação M19 e M21;
- **Próximo a porta:** Sensor de movimentação M020 e um sensor de temperatura T003;
- **Próximo ao closet:** Sensores de movimentação M18 e M22.

Closets: Sem sensores instalados.

Banheiro:

- **Porta:** Sensor de movimentação M026;
- **Pia:** Sem sensores;
- **Zona de passagem no banheiro:** Sensor de movimentação M027 e sensor de temperatura T005;
- **Área do sanitário:** Sensor de movimentação M025.

Atividades dos Usuários:

Para o dataset mais recente (WSU Tulum Smart Apartment 2009-2010 Two Resident Testbed) os autores conseguiram identificar 16 atividades diferentes através dos sensores utilizados, entretanto, verificando o *dataset*, somente 14 atividades foram encontradas:

- Dormindo na cama;
- Higiene pessoal;
- Tomando banho;
- Saindo de casa;
- Entrando em casa;
- Preparando refeição;
- Assistindo TV;
- Comendo;
- Transição da cama para o banheiro;
- Mesa de trabalho;
- Yoga;
- Quarto de trabalho 1;
- Quarto de trabalho 2;
- Sala de estar do trabalho.

Já nos outro *dataset* foram mapeadas apenas 10 atividades por usuário, entretanto, verificando o *dataset*, somente 7 atividades foram encontradas:

- Fazendo café da manhã;
- Fazendo almoço;
- Saindo de casa;
- Entrando em casa;
- Assistindo TV;
- Comendo o café da manhã;
- Comendo um lanche.

As atividades Saindo de casa, Entrando em casa e Assistindo TV são comuns aos dois *datasets*. No primeiro *dataset* somente a primeira atividade está diretamente ligada à um usuário, já no segundo *dataset* somente as duas últimas atividades estão mapeadas para algum usuário. Os autores não vinculam as outras atividades aos usuários, ou seja, não dá para saber quem foi que realizou aquela atividade. Porém, como algumas atividades acontecem durante um período que somente um usuário foi mencionado, acredita-se que essas atividades estejam relacionada aquele usuário.

Estrutura do Arquivo de Dados: O arquivo é dividido em 5 colunas:

- **Data de sensoriamento:** Informa o período em que os dados foram coletados de cada sensor;
- **Tempo:** Em que horário aquele sensor foi utilizado para captar os dados do ambiente;
- **Sensor:** Identifica qual sensor da casa foi utilizado para capturar aquela informação (ex., M046);
- **Informação do sensor:** Informa se o sensor está ligado ou não, além de algumas outras informações que variam dependendo do tipo do sensor;
- **Tags:** Denotando em que momento começa determinada atividade do usuário 1 ou 2.

3. WSU Aruba Smart Apartment - One Resident Testbed

Descrição: Um *dataset* com atividades diárias de uma mulher em sua residência, que recebe visitas regulares de seus filhos e netos. As atividades são anotadas, com marcações mostrando o início e o fim da atividade.

Dataset:

- WSU Aruba Smart Apartment 2009 Two Resident Testbed (Dataset 17)

Categorização dos sensores: Os autores categorizam os sensores em três tipos diferentes.

- **M:** Sensores de movimento;
- **D:** Sensores de porta;
- **T:** Sensores de temperatura.

Formato da casa:

- Porta de entrada;
- Sala de estar;
- Cozinha;
- Corredor da sala para os quartos;
- Banheiro social;
- Porta para a garagem;
- Escritório;
- Quarto 1;
- Quarto 2;
- Banheiro do Quarto 2;
- Closet do Quarto 2;

Disposição dos sensores no ambiente:**Porta de Entrada:**

- Sensor de porta D001.

Sala de estar:

- **Lado esquerdo:** Sensor de temperatura T002;
- **Meio:** Sensores de movimentação M012 - M014;
- **Lado direito:** Dois sensores de movimentação M009 e M010.

Cozinha:

- **Lado esquerdo:** Sensor de movimentação M016 e sensor de temperatura T003;
- **Meio:** Sensor de movimentação M019;
- **Entrada da Cozinha até a porta de traz:** Sensores de movimentação M015 e M017, além disso, sensor de porta D002.

Corredor da sala para os quartos:

- Sensores de movimentação M021 e M022;
- Sensor de temperatura T004.

Banheiro social:

- **Da porta de entrada pelo corredor:** Sensor de porta D003 e sensor de movimentação M031;
- **Porta de entrada 2:** Sensor de movimentação M029.

Porta para a garagem: Sensor de porta D004.**Escritório:**

- **Entrada:** Sensores de movimentação M028;
- **Meio:** Sensor de movimentação M027;
- **Canto superior:** Sensor de movimentação M026;

- **Canto inferior:** Sensor de movimentação M025 e sensor de temperatura T005.

Quarto 1:

- **Entrada:** Sensores de movimentação M023;
- **Meio:** Sensor de movimentação M024.

Quarto 2:

- **Porta de Entrada:** Sensores de movimentação M008;
- **Porta de entrada até porta do banheiro:** Sensor de movimentação M005 e M004;
- **Meio para a esquerda:** Sensor de temperatura T001 e sensor de movimentação M003;
- **Meio:** Sensor de movimentação M007;
- **Canto inferior esquerdo:** Sensor de movimentação M002;
- **Canto inferior direito:** Sensor de movimentação M001.

Closet do Quarto 2: Sem sensores instalados.

Banheiro do Quarto 2: Sem sensores instalados.

Atividades dos Usuários:

Os autores conseguiram identificar 11 atividades diferentes através dos sensores utilizados, além disso, eles também informam quantas vezes cada atividade foi realizada:

- Dormindo (401);
- Saindo de casa (431);
- Entrando em casa (431);
- Preparando refeição (1606);
- Comendo (257);
- Transição da cama para o banheiro (157);
- Descansando (2910);
- Trabalho (171);
- Lavando os pratos (65);
- Limpando a casa (33);
- *Resperate* (6).

Estrutura do Arquivo de Dados: O arquivo é dividido em 5 colunas:

- **Data de sensoriamento:** Informa o período em que os dados foram coletados de cada sensor;
- **Tempo:** Em que horário aquele sensor foi utilizado para captar os dados do ambiente;
- **Sensor:** Identifica qual sensor da casa foi utilizado para capturar aquela informação (ex., M046);

- **Informação do sensor:** Informa se o sensor está ligado ou não, além de algumas outras informações que variam dependendo do tipo do sensor;
- **Tags:** Denotando em que momento começa determinada atividade do usuário.

4. HH Smart Apartment - One Resident Testbed

Descrição: Um *dataset* com atividades diárias de uma pessoa adulta em sua residência. As atividades do usuário são parcialmente anotadas, com marcações mostrando o início e o fim da atividade.

Dataset:

- HH129 Single-resident apartment data (Dataset 51)

Existem vários *datasets* do conjunto HH que seguem a mesma estrutura do *dataset* escolhido.

Categorização dos sensores: Os autores categorizam os sensores em sete tipos diferentes:

- **M:** Sensores de movimentação infravermelho;
- **D:** Sensores de porta magnéticos;
- **T:** Sensores de temperatura;
- **BA:** Sensores de bateria;
- **L & LL:** Sensores de interruptores;
- **LS:** Sensores de luz;
- **MA:** Sensores de movimento infravermelho em áreas grande.

Formato da casa: O apartamento possui uma cozinha, dois quartos e um banheiro.

Disposição dos sensores no ambiente:

Cozinha:

- **Porta de entrada:** Um sensor de movimentação M001 e um sensor de porta D002;
- **Ao lado do fogão:** Um sensor de porta D001 e um sensor de temperatura T101;
- **Parede que faz divisa com o quarto 1:** Sensores de movimentação M003 - M004.

Banheiro: Sensor de movimentação M012.

Quarto 1:

- **Entrada:** Sensor de movimentação M005;
- **Parede que faz divisa com a cozinha:** Sensores de movimentação M006 e M007;

- **Parede que faz divisa com o outro quarto:** Sensores de movimentação M010 e M011;
- **Parede a direita:** Sensores de movimentação M008 e M009.

Quarto 2: Não possui sensores de monitoramento.

Observando a planta do apartamento com a disposição dos sensores no ambiente, pode-se perceber que nem todos os sensores listados pelos autores são mostrados na planta. Portanto, assume-se que a versão final da planta com todos os sensores no ambiente não foi disponibilizada pelos autores.

Atividades dos Usuários:

Os autores conseguiram identificar 13 atividades diferentes através dos sensores utilizados (algumas atividades possuem subatividades, por exemplo, Comendo, têm comendo café da manhã, comendo almoço e comendo o jantar):

- Dormindo;
- Saindo de casa;
- Entrando em casa;
- Preparando refeição;
- Preparando café da manhã;
- Preparando almoço;
- Preparando jantar;
- Comendo;
- Comendo café da manhã;
- Comendo almoço;
- Comendo jantar;
- Transição da cama para o banheiro;
- Descansando;
- Trabalhando na mesa;
- Trabalhando na escrivaninha;
- Banheiro;
- Lavando os pratos;
- Lavando os pratos do café da manhã;
- Lavando os pratos do almoço;
- Lavando os pratos do jantar;
- Higiene pessoal;
- Assistindo TV;
- Limpando a casa (33);

Estrutura do Arquivo de Dados: O arquivo é dividido em 5 colunas:

- **Data de sensoriamento:** Informa o período em que os dados foram coletados de cada sensor;
- **Tempo:** Em que horário aquele sensor foi utilizado para captar os dados do ambiente;

- **Sensor:** Identifica qual sensor da casa foi utilizado para capturar aquela informação (ex., M046);
- **Informação do sensor:** Informa se o sensor está ligado ou não, além de algumas outras informações que variam dependendo do tipo do sensor;
- **Tags:** Denotando em que momento começa determinada atividade do usuário.

5. Conclusão

Todos os *datasets* analisados possuem dados de atividades diárias de moradores em suas residências. Observando os sensores utilizados, é possível perceber que sensores de movimentação e temperatura são utilizados em todos os *datasets*. O sensor de porta não foi utilizado somente nos *datasets* de Tulum.

Utilizando os sensores citados é possível saber quando o usuário entra e sai de casa, a movimentação do usuário dentro da casa, sabendo em quais momentos ele está em cada cômodo e por quanto tempo, e a temperatura da casa. Com relação às atividades, existem pequenas variações de um *dataset* para o outro, por exemplo, um *dataset* não vai ter a atividade “Assistindo TV” enquanto todos os outros vão possuir essa atividade. Entretanto, existem 4 atividades que se repetem em todos os *datasets*, são elas: i) Transição da cama para o banheiro; ii) Higiene pessoal; iii) Dormindo; e iv) Trabalho. No arquivo dos *datasets* essas atividades possuem anotações informando o começo e o fim da atividade, entretanto, existem alguns problemas com as anotações.

Nos *datasets* de Kyoto e Tulum somente algumas atividades são anotadas para os usuários, assim, existem algumas atividades que não dá para saber se foram realizadas pelo usuário 1 ou 2. Em HH129 existe somente um usuário na residência, o problema é que os dados são parcialmente anotados, então existem atividades que não foram mapeadas, além de poder existir, por exemplo, atividades que só possuem a marcação de começo, faltando a marcação de finalização da atividade.

O único *dataset* que não possui esses problemas é o de Aruba, pois existe somente um usuário na residência e o *dataset* é todo anotado, com marcações mostrando o começo e o final de cada atividade realizada pelo usuário. Portanto, conclui-se que o melhor *dataset* para a realização dos testes da abordagem proposta seja o *dataset* de Aruba.